

Europäisches Patentamt

European Patent Office

Office européen des brevets



EP 0 773 290 A2 (11)

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication: 14.05.1997 Bulletin 1997/20 (51) Int. Cl.⁶: C12N 15/12, C07K 14/705

(21) Application number: 96116108.0

(22) Date of filing: 08.10.1996

(84) Designated Contracting States: AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC **NL PT SE**

(30) Priority: 09.10.1995 JP 261440/95 24.04.1996 JP 102451/96

(71) Applicant: KOWA COMPANY, LTD. Naka-ku Nagoya-shi Aichi-ken (JP)

(72) Inventors:

- · Saito, Yasushi Chiba-shi, Chiba (JP)
- · Iwasaki, Akio Tsuchiura-shi, Ibaraki (JP)

· Arai, Koichi Urawa-shi, Saitama (JP)

· Yamazaki, Hiroyuki Higashimurayama-shi, Tokyo (JP)

(74) Representative: Wächtershäuser, Günter, Prof. Dr.

Patentanwalt,

Tal 29

80331 München (DE)

Remarks:

The applicant has subsequently filed a sequence listing and declared, that it includes no new matter.

(54)Novel LDL receptor analog protein and the gene coding therefor

The present invention is drawn to the gene of a novel LDL receptor family receptor which participates in lipoprotein metabolism, a critical factor that triggers the onset of arteriosclerosis.

The invention provides DNA having a nucleotide sequence as shown by Sequence ID No. 1 or No.5 is disclosed as well as rabbit tissue or human tissue LDL receptor analog protein having an amino acid sequence of Sequence ID No. 2 or 6 coded by such DNA.

Description

Background of the Invention:

Field of the Invention

The present invention relates to a novel LDL receptor analog protein having a structure similar to that of LDL receptors that are responsible for the homeostasis mechanism of intracellular cholesterol and extensively participates in serum lipid metabolism, which is a critical factor that triggers the onset of arteriosclerosis. The invention also relates to the gene coding for the protein.

2) Description of the Related Art

Abnormality in serum lipid metabolism is one of the most critical risk factors in the onset and progress of arteriosclerosis. Serum lipids, together with apolipoproteins, are transformed into lipoproteins primarily in the liver, secreted therefrom, transported by blood, and taken up by a variety of tissue cells.

Uptake of lipoproteins into cells occurs primarily by the mediation of receptors of respective lipoproteins. It is known that low density lipoproteins (LDL), which are taken into cells by specific membrane receptors, called LDL receptors, are metabolized within the cells and utilized as cell membrane components or similar substances. Detailed analysis of familial hyperchlolesterolemia, which is a genetic disease accompanied by notable hyperchlolesterolemia due to abnormality of LDL receptors, has clarified details of the mechanism of homeostasis achieved by LDL receptors with respect to intracellular cholesterol.

It has been suggested that living bodies have not only LDL receptors but also cell membrane receptors that recognize other lipoproteins. From analyses of WHHL rabbits, which are model animals lacking LDL receptors, it was found that receptors which takes principally apo-E-containing lipoproteins as ligands (remnant receptors) are present in the liver. It is also predicted that there may be HDL receptors whose ligands are high density lipoprotein (HDL). However, to date, details of the structures and functions of these receptors have not yet been elucidated. It has also been known that foaming of macrophages plays an active role in the formation of atherosclerosis, is deeply participated. Macrophages foam by taking up modified LDL —not normal LDL— which have undergone oxidation, acetylation, or glycation. There have recently been discovered receptors to modified LDL which are called scavenger receptors. The scavenger receptors have been identified to be membrane receptors that have a structure completely different from that of LDL receptors.

Recent research using molecular biological techniques has identified the genes of LRP (LDL receptor-associated protein), gp 330, and VLDL receptors. The receptors have been found to have structures very similar to those of LDL receptors. From analyses of these receptors, it is believed that a plurality of lipoprotein receptors are present in living bodies, and that they are closely related to lipid metabolism. LDL receptors studied in detail by Brown and Goldstein [Brown, M.S. and Goldstein, J.L. (1986) Science 232, 34-47] are known to play an important role in the homeostasis of lipoprotein metabolism in vivo, recognizing apo-B-100 and apo-E and taking primarily LDL as their ligands. Also, LRP, which is a macroprotein, has been found to primarily recognize apo-E and to take β -VLDL or chylomicron remnant as a ligand. Moreover, it has been recently reported that LRP takes an α_2 -macroglobulin/protease complex or a plasminogen activator/plasminogen activator inhibitor-1 complex as a ligand, and that LRP is a protein identical to the ap-macroglobulin receptor. When these findings are taken together, LRP is considered to have a wide variety of functions in living bodies [Herz, J., Hamann, U., Rogne, S., Myklebost, O., Gausepohl, H. and Stanley, K.K. (1989) EMBO J. 7(13), 4119-4127; Brown, M.S., Herz, J., Kowal, R.C. and Goldstein, J.L. (1991) Current Opinion in Lipidology 2, 65-72; Herz, J. (1993) Current Opinion in Lipidology 4, 107-113]. The gp 330, which was first identified as an antigen inducing rat Heymann nephritis, has been reported to have a ligand-binding capacity similar to that possessed by CRP ag-macroglobulin receptor [Raychowdhury, R., Niles, J.L., McCluskey, R.T. and Smith, J.A. (1989) Science 244, 1163-1165; Pietromonaco, S., Kerjaschki, D., Binder, S., Ullrich, R. and Farquhar, G. (1990) Proc. Natl. Acad. Sci. U.S.A. 87, 1811-1815]. In addition, recently discovered VLDL receptors, which are found to take VLDL as a ligand, are considered to have new functions including fatty acid metabolism, because they are predominantly found in tissues of the heart and muscles though they are rarely found in the liver [Takahashi, S., Kawarabayashi, Y., Nakai, T., Sakai, J. and Yamamoto, T. (1992) Proc. Natl. Acad. Sci. USA 89, 9252-9256].

Functions of these newly found receptors as lipoprotein receptors have been gradually elucidated through detailed in vitro analyses. However, significance of respective receptors in living bodies has mostly been left unknown. In addition, relations to remnant receptors, HDL receptors, etc., which have conventionally been identified or suggested by biochemical techniques, remain unknown. Presently, it is considered that these newly found receptors are products of genes different from those of the latter receptors. Thus, more lipoprotein receptors than originally guessed have become considered to participate in lipoprotein uptake into cells while interacting with each other to thereby function to maintain homeostasis of lipid metabolism in living bodies. However, from structural analyses of the genes of the afore-

mentioned newly-identified receptors, it is predicted that the genes of these receptors that take lipoproteins as ligands are developed from the same gene from which LDL receptors was developed, and thus they are within the same genetic family. This suggests that lipoprotein receptors that have conventionally been proposed may have structures similar to those of LDL receptors.

Accordingly, an object of the present invention is to provide the gene of a novel receptor in the LDL receptor family, as well as a protein coded by the gene.

The present inventors conducted careful studies so as to attain the above object, and found that by using part of rabbit LDL receptor cDNA as a probe there can be obtained a DNA fragment coding for a peptide having a structure similar to that of LDL receptors. Moreover, when using part of the obtained cDNA as a probe, a cDNA fragment having a sequence similar to that of the cDNA can be obtained from the human tissue cDNA library. The present invention was accomplished based on these findings.

Summary of the Invention

5

15

50

The present invention provides DNA having a nucleotide sequence shown by Sequence ID No. 1 or No. 5; an LDL receptor analog protein having an amino acid sequence coded by the DNA; a recombinant vector comprising the DNA and a replicable vector; transformant cells which harbor the recombinant vector; and a method for the production of the LDL receptor analog protein.

Description of Preferred Embodiment

The cDNA of the present invention may be prepared, for example, by the following process.

Briefly, the process includes the following steps. (1) Through the use of rabbit LDL receptor cDNA as a probe, positive clones are screened out of a rabbit liver cDNA library. (2) Recombinant DNA is prepared using the separated positive clones, and a cDNA fragment is cut out of the resultant recombinant DNA through a treatment using a restriction enzyme. The cDNA fragment is integrated into a plasmid vector. (3) Host cells are transformed using the obtained cDNA recombinant vector to thereby obtain transformant cells of the present invention. The obtained transformant cells are incubated so as to obtain a recombinant vector containing a DNA fragment of the present invention. The nucleotide sequence of the DNA fragment of the present invention contained in the resultant recombinant vector is determined. (4) In tissue of a living body, there is detected expression of mRNA indicated by the nucleotide sequence of the cDNA of the present invention by using RNA blot hybridization method. (5) Through use of a rabbit cDNA fragment as a probe, positive clones are screened out of a human tissue cDNA library, and the nucleotide sequence of the clones is determined. (6) A recombinant vector for expression is prepared using the cDNA of the present invention. Through use of the thus-obtained vector, host cells are transformed to thereby obtain the transformants of the present invention. (7) Ligands that are bound to protein expressed by the obtained transformants are detected by ligand blotting.

Each of the above-described steps will next be described.

(1) Screening for positive clones from a rabbit liver cDNA library:

A cDNA library may be prepared by the use of mRNA obtained from rabbit liver, reverse transcriptase, and a suitable vector, e.g., commercially available \(\lambda \text{gt10} \) vector.

A cDNA library thus prepared using λgt10 as a vector is subjected to a screening for positive clones by the application of a DNA hybridization method employing a cDNA probe, to thereby separate positive clones [Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) In: Molecular Cloning: A Laboratory Manual, pp 9.47-9.58, Cold Spring Harbor Laboratory Press].

An exemplary cDNA which may be used as a probe is rabbit LDL receptor cDNA. Positive clones may be detected by autoradiography employing a DNA probe labelled with a radioisotope (³²P).

(2) Preparation of a cDNA recombinant vector:

Recombinant vector λ gt10 phage DNA is extracted from the isolated positive clones and purified. The resultant purified recombinant vector λ gt10 phage DNA is digested with a restriction enzyme EcoRI, to thereby separate a cDNA fragment from the vector DNA. The obtained cDNA fragment is integrated with a plasmid vector for cloning that has been similarly digested with EcoRI, thereby obtaining a recombinant plasmid vector. An exemplary plasmid vector which may be used is pBluescript II.

(3) Recombinant vector, transformation of host cells using the recombinant vector, and preparation of DNA:

The obtained cDNA recombinant vector is introduced into a variety of host cells that are capable of utilizing the

genetic marker possessed by the recombinant vector, to thereby transform the host cells. Host cells are not particularly limited, with *E. coli* being preferred. For example, a variety of variants of the *E. coli* K12 strain, e.g., HB-101, may be used. In order to introduce the recombinant vector into host cells, a competent cell method may be used in combination with a treatment with calcium.

The thus-obtained transformant cells are cultured in a selective medium in accordance with the genetic marker of the vector. The recombinant vector of the present invention is collected from the cultured cells. The DNA nucleotide sequence of the cDNA contained in the obtained recombinant vector can be determined through use of a dideoxy sequence method [Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acd. Sci. USA 74, 5463-5467].

(4) RNA blot hybridization:

5

30

40

50

55

The expression in tissue of mRNA, indicated by the nucleotide sequence of the cDNA of the present invention, is detected using RNA blot hybridization.

First, mRNA is prepared using rabbit tissue. Commercially available oligo(dT)cellulose column may be used for the preparation. In order to prepare mRNA from human tissue, there may be used a commercially available nylon membrane on which tissue poly(A)*RNA from a variety of sources is present.

An exemplary probe is the rabbit cDNA obtained in the above-described step (3). mRNA may be detected by autoradiography employing a DNA probe labelled with a radioisotope (32P).

(5) Screening of human tissue cDNA library for positive clones, and determination of nucleotide sequence:

An exemplary human tissue cDNA library which may be used is a commercially available human brain cDNA library.

Screening and nucleotide sequencing of the human brain cDNA library may be performed using a fragment of rabbit cDNA of the present invention as a probe in a manner similar to that used for the aforementioned rabbit liver cDNA library.

(6) Preparation of a recombinant vector for expression and transformation of host cells using the recombinant vector for expression:

In order to prepare an LDL receptor analog protein through use of cDNA of the present invention, the obtained cDNA and a vector for expression are first bonded to each other to thereby create a recombinant vector for expression. Vectors for expression which may be used for bonding are not particularly limited. For example, pBK-CMV may be used.

Host cells are transformed using the thus-obtained recombinant vector for expression, to thereby obtain a transformant cell of the present invention. The obtained transformant cell is cultured so as to obtain cells that are capable of expressing the protein of the invention. Host cells are not particularly limited. For example, CHO cells may be used. In order to introduce the recombinant vector for expression into host cells, a calcium phosphate method may be used.

The thus-prepared transformant cells are incubated in a selective medium in accordance with the genetic marker of the vector, so as to express the LDL receptor analog protein of the present invention.

(7) Ligand analysis of the protein by ligand blotting:

After the resultant transformant cells are incubated, the expressed LDL receptor analog protein is solubilized using a solubilizer, e.g., Triton X-100, to thereby obtain a membrane protein fraction. The fraction is separated using SDS-PAGE, and transferred onto, for example, a nitrocellulose membrane. Using a radio-labelled (125 I) lipoprotein as a probe, the analog protein can be detected by autoradiography. Exemplary lipoproteins which may be used include β -VLDL and LDL.

Examples:

The present invention will next be described in detail by way of example, which should not be construed as limiting the invention.

Example 1:

Preparation of a rabbit liver cDNA library:

From tissue of the liver of a male Japanese white rabbit, intact RNA was extracted through a guanidium thiocy-anate/cesium chloride method. The obtained intact RNA was subjected to an oligo (dT) cellulose column method to

thereby obtain purified poly(A)+RNA.

cDNA was synthesized in accordance with a method of Gubler and Hoffman [Gubler, U. and Hoffman, B.J. (1983) Gene 25, 263]. Briefly, cDNA was synthesized employing rabbit liver poly(A)⁺RNA (as a template), a random primer, and moloney murine leukemia virus reverse transcriptase. The synthesized cDNA was transformed into double-stranded DNA using DNA polymerase I, and then subjected to an EcoRI methylase treatment. By the use of T4 DNA polymerase, the DNA was blunt-ended. The blunt-ended DNA was ligated to phosphorylated EcoRI linker pd (CCGAATTCGG) using a T4 DNA ligase, and the resultant ligated product was subjected to an additional digestion with EcoRI. cDNA fragments having a size not less than 1 kb were selected by agarose gel electrophoresis, and integrated into the EcoRI-digested site of λgt10 phage DNA using a T4 DNA ligase. The phage DNA was packaged *in vitro*, to thereby establish a rabbit liver cDNA library.

Example 2:

15

Cloning of cDNA of receptors in the rabbit LDL receptor family:

The cDNA library (1,000,000 plaques) prepared in Example 1 was subjected to screening using a plaque hybridization method and employing as a probe a segment of the cDNA obtained from a ligand binding region, the functional region, of the rabbit LDL receptor. Hybridization was performed at 42°C using 5 x SSC, 30% formamide, 1% SDS, 5 x Denhardt's, and 100 μg/ml salmon sperm DNA (ssDNA), followed by washing with 0.3 x SSC/0.1% SDS at 48°C. As a result, several positive clones were obtained. These cDNA clones were separated by performing this plaque hybridization method in a plurality of times. Subsequently, a cDNA fragment of each phage was subcloned into a plasmid vector pBluescript II, and the nucleotide sequence was analyzed using a dideoxy sequence method [Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acd. Sci. USA 74, 5463-5467]. Based on a putative amino acid sequence, LDL receptors themselves were excluded, and cDNA clones having a sequence very similar to that of LDL receptors were identified. Using these clones as cDNA probes, the cDNA library was screened to thereby obtain overlapping two clones. These were employed as new probes and similar procedure was performed, so as to obtain 5 cDNA clones. The DNA nucleotide sequence determined by these cDNA clones are shown as Sequence ID No. 3. The total length of the sequence was 6961 bp. In the open reading frame of 6639 bp (Sequence ID No. 1) which contained a sequence exhibiting high homology with LDL receptors, there existed on the 5' side an ATG codon which was presumably a translation initiating site and a successive highly hydrophobic sequence consisting of about 30 amino acids. Accordingly, the obtained cDNA was considered to contain the entirety of its length. A putative amino acid sequence is shown as Sequence ID No. 2. The protein consisted of 2213 amino acids. Comparison of the amino acid sequence of the protein with other amino acid sequence data registered at the Genebank, there was a very high similarity to LDL receptors. That is, amino acids 700 - 1,100 in the sequence were very similar to the EGF precursor homology region of LDL receptors, and amino acids 1,100 - 1,640 were also very similar to the ligand binding region of LDL receptors. When the amino acid sequence of the subject protein was compared with other lipoprotein receptor LRP, gp330, and VLDL receptors, similarity was not as high as that observed for LDL receptors. On the C-terminal side of the amino acid sequence of the protein, there was found a highly hydrophobic region which was very similar to the transmembrane region of LDL receptors.

Example 3:

40

From liver tissue and brain tissue of a male Japanese white rabbit, intact RNA was extracted through a guanidium thiocyanate/cesium chloride method. The obtained intact RNA was subjected to an oligo (dT) cellulose column method to thereby obtain purified poly(A) $^+$ RNA. The poly(A) $^+$ RNA specimens (10 μ g each) was modified via a glyoxal method, electrophoresed on 1% agarose gel, and transferred onto a nylon membrane.

For human tissue mRNA, commercially available nylon membranes blotted with human tissue poly(A)*RNA from various sources were used.

Using as a probe part of a 32 P-labelled rabbit cDNA of the present invention, hybridization was performed at 42 °C using 50% (rabbit) or 40% (human) formamide, 0.1% SDS, 50 mM phosphate buffer, 5 x Denhardt's, 5 x SSC, and 200 μ g/ml of ssDNA, followed by washing with 0.1 x SSC and 0.1% SDS at 50°C. Autoradiography was performed at $^{-70}$ °C for 2 days in the presence of intensifying screen. As a result, in both rabbit liver tissue and brain tissue, mRNA of about 7 kb was detected as well as mRNA of about 15 kb which was considered to result from alternative splicing or polyadenylation. The size of the mRNA of about 7 kb coincided with that of the rabbit cDNA of the present invention. Also, in human liver tissue and brain tissue, it was confirmed that mRNA having the same size was expressed.

Example 4:

Screening of human brain cDNA library for positive clones and determination of the nucleotide sequence of cDNA frag-

The human brain cDNA library used in this Example was a commercially obtained cDNA library which was constructed using \(\lambda\)gt10 as a vector. Using partial cDNA of the present invention as a probe, screening of the cDNA library (300,000 plaques) was performed using a plaque hybridization method. Procedures of screening, cloning, and sequencing were as described in Example 2 of the present invention.

As a result of screening of the human brain cDNA library, positive clones containing a DNA fragment of about 3 kb were obtained. Analysis of the nucleotide sequence of part of the cDNA fragment revealed that the fragment was highly homologous to the cDNA of the present invention (Sequence ID No. 4).

Example 5:

Cloning of cDNA of receptors in the human LDL receptor family:

A human brain cDNA library was subjected to screening using fragments of the cDNA of the present invention and fragments of the cDNA obtained in Example 4 as probes. Procedures of screening, cloning, and sequencing were as described in Example 2 of the present invention.

Through screening of the human brain cDNA library, two positive clones containing cDNA fragments of about 6 kb and about 3 kb were obtained. When their nucleotide sequence was analyzed, they were identified to be a cDNA clone containing the cDNA nucleotide sequence obtained in Example 4 and a cDNA clone that overlapped therewith. Using part of these cDNAs as probes, procedures similar to those as described above were performed, to thereby obtain another cDNA clone. The DNA nucleotide sequence indicated by these cDNA clones are shown as Sequence ID No. 7. The total length of the sequence was 6,843 bp. There was an open reading frame having a size of 6,642 bp (Sequence ID No. 5). A putative amino acid sequence is shown as Sequence ID No. 6. The protein consisted of 2,214 amino acids. Comparison of the amino acid sequence with that of rabbit protein shown by Sequence ID No. 2 revealed high homology of not less than 94%.

Example 6:

Creation of cells that express receptors in the rabbit LDL receptor family:

The cDNA as shown by Sequence ID No. 3 was ligated to phosphorylated EcoRI linker pd (CCGAATTCGG) by the use of a T4 DNA ligase, and the resultant ligated product was digested with EcoRI. Separately, a vector for expression, pBK-CMV was digested with EcoRI. The aforementioned DNA was ligated to the EcoRI-digested site of the vector using a T4 DNA ligase.

Using the resultant recombinant expression vector in a calcium phosphate method [Chen, C. and H. Okayama (1987) Mol. Cell. Biol. 7, 2745-2752], host cells (CHO-1d1A7) were transformed. The resultant transformants were incubated in a Ham's F-12 selective medium supplemented with 500 µg/ml of G418, and viable cells were separated as LDL receptor analog protein-expressing cells. The cells were incubated further in the aforementioned medium.

Example 7:

Ligand analysis of the LDL receptor analog protein by ligand blotting:

The obtained LDL receptor analog protein-expressing cells and control cells were suspended in a buffer solution containing 200 mM Tris-maleic acid (pH 6.5), 2 mM calcium chloride, 0.5 mM PMSF, 2.5 µM leupeptin, and 1% Triton X-100, to thereby solubilize the membrane protein. Solubilized membrane protein fractions were obtained through centrifugation, and electrophoresed by a 4.5-18% gradient SDS-PAGE. Thereafter, the protein was transferred onto a nitrocellulose membrane.

Incubation was performed in a buffer of 50 mM Tris-HCl (pH 8.0) containing ¹²⁵I-labelled β-VLDL (10 μg/ml), 2 mM calcium chloride, and 5% bovine serum albumin. Autoradiography was performed at room temperature.

A single band of about 250 kDa was detected in membrane protein fractions prepared using the present proteinexpressing cells. This size coincided well with the molecular weight of 248 kDa calculated regarding the amino acid sequence (Sequence ID No. 2) deduced from the cDNA of the present invention. Although a similar band was detected for control cells, the expression level was much lower as compared with the case of the present protein-expressing cells.

Since the protein coded by the cDNA of the present invention is considered to be a novel LDL receptor family recep-

6

5

10

15

35

30

45

tor, it is expected that through analyses of this protein, details of lipoprotein metabolism mediated by the membrane receptor will be elucidated, and pathology of abnormal lipid metabolism which triggers onset and progress of arteriosclerosis will be clarified.

Sequence ID No. 1

Length of the Sequence: 6639

Type: nucleic acid

Strandedness: double

Topology: linear

Molecular type: cDNA to mRNA

Sequence:

5

10

15

20

25

30

40

45

50

ATGGCGACAC	GGAGCAGCAG	GAGGGAGTCG	CGACTCCCCT	TCCTATTCAC	CCTGGTCGCG	60
CTGCTGCCGC	CCGGGGCTCT	CTGCGAGGTG	TGGACGCGGA	CACTGCACGG	CGGCCGCGCG	120
CCCTTACCCC	AGGAGCGGGG	CTTCCGCGTG	GTGCAGGGCG	ACCCGCGCGA	GCTGCGGCTG	180
TGGGAGCGCG	GGGATGCCAG	GGGGGCGAGC	CGGGCGGACG	AGAAGCCGCT	CCGGAGGAGA	240
CGGAGCGCTG	CCCTGCAGCC	CGAGCCCATC	AAGGTGTACG	GACAGGTCAG	CCTCAATGAT	300
TCCCACAATC	AGATGGTGGT	GCACTGGGCC	GGAGAGAAAA	GCAACGTGAT	CGTGGCCTTG	360
GCCCGGGACA	GCCTGGCGTT	GGCCAGGCCC	AGGAGCAGTG	ATGTGTACGT	GTCTTATGAC	420
TATGGAAAAT	CATTCAATAA	GATTTCAGAG	AAATTGAACT	TCGGCGCGGG	AAATAACACA	480
GAGGCTGTGG	TGGCCCAGTT	CTACCACAGC	CCTGCGGACA	ACAAACGGTA	CATCTTCGCA	540
GATGCCTACG	CCCAGTATCT	CTGGATCACG	TTTGACTTCT	GCAACACCAT	CCATGGCTTT	600
TCCATCCCGT	TCCGGGCAGC	TGATCTCCTA	CTCCACAGTA	AGGCCTCCAA	CCTTCTCCTG	660
GGCTTCGACA	GGTCTCACCC	CAACAAGCAG	CTGTGGAAGT	CGGATGATTT	TGGCCAGACC	720
TGGATCATGA	TTCAAGAACA	CGTGAAGTCC	TTTTCTTGGG	GAATTGATCC	CTATGACAAA	780
CCAAACACCA	TCTACATCGA	ACGGCACGAA	CCTTCTGGCT	ACTCCACGGT	TTTCCGAAGT	840
ACAGACTTCT	TCCAGTCCCG	GGAAAACCAG	GAAGTGATCT	TGGAGGAAGT	GAGAGACTTT	900
CAGCTTCGGG	ACAAGTACAT	GTTTGCTACA	AAGGTGGTGC	ATCTCTTGGG	CAGTCCACTG	960
CAGTCTTCTG	TCCAGCTCTG	GGTCTCCTTT	GGCCGGAAGC	CCATGCGGGC	CGCCCAGTTT	1020
GTTACAAGAC	ATCCTATCAA	CGAATATTAC	ATCGCGGATG	CCTCGGAGGA	CCAGGTGTTT	1080
GTGTGTGTCA	GTCACAGCAA	CAACCGCACC	AACCTCTACA	TCTCGGAGGC	AGAGGGCTTG	1140
AAGTTCTCTC	TGTCCCTGGA	GAACGTGCTC	TACTACACCC	CGGGAGGGGC	CGGCAGTGAC	1200
ACCTTGGTGA	GGTACTTTGC	AAATGAACCG	TTTGCTGACT	TCCATCGTGT	GGAAGGGTTG	1260

	CAGGGAGTCT	ACATTGCTAC	TCTGATTAAT	GGTTCTATGA	ATGAGGAGAA	CATGAGATCT	1320
.	GTCATCACCT	TTGACAAAGG	GGGCACCTGG	GAATTTCTGC	AGGCTCCAGC	CTTCACGGGG ·	1380
,	TATGGAGAGA	AAATCAACTG	TGAGCTGTCC	GAGGGCTGTT	CCCTCCACCT	GGCCCAGCGC	1440
	CTCAGCCAGC	TGCTCAACCT	CCAGCTCCGG	AGGATGCCĊA	TCCTGTCCAA	GGAGTCGGCG	1500
10	CCTGGCCTCA	TCATTGCCAC	GGGCTCAGTG	GGAAAGAACT	TGGCTAGCAA	GACAAACGTG	1560
	TACATCTCTA	GCAGTGCTGG	AGCCAGGTGG	CGAGAGGCAC	TTCCTGGACC	TCACTACTAT	1620
	ACATGGGGAG	ACCATGGCGG	CATCATCATG	GCCATTGCCC	AAGGCATGGA	AACCAACGAA	1680
15	CTGAAGTACA	GTACCAACGA	AGGGGAGACC	TGGAAAGCCT	TCACCTTCTC	TGAGAAGCCC	1740
	GTGTTTGTGT	ATGGGCTCCT	CACGGAACCC	GGCGAGAAGA	GCACGGTCTT	CACCATCTTT	1800
20	GGCTCCAACA	AGGAGAACGT	GCACAGCTGG	CTCATCCTCC	AGGTCAATGC	CACAGACGCC	1860
20	CTGGGGGTTC	CTTGCACAGA	GAACGACTAC	AAGCTCTGGT	CACCATCTGA	TGAGCGGGGG	1920
	AATGAGTGTT	TGCTTGGACA	CAAGACTGTT	TTCAAACGGA	GGACCCCGCA	CGCCACATGC	1980
?5	TTTAACGGAG	AAGACTTTGA	CAGGCCGGTG	GTTGTGTCCA	ACTGCTCCTG	CACCCGGGAG	2040
	GACTATGAGT	GTGACTTTGG	CTTCCGGATG	AGTGAAGACT	TGGCATTAGA	GGTGTGTGTT	2100
•	CCAGATCCAG	GATTTTCTGG	AAAGTCCTCC	CCTCCAGTGC	CTTGTCCCGT	GGGCTCTACG	2160
30	TACAGGCGAT	CAAGAGGCTA	CCGGAAGATT	TCTGGGGACA	CCTGTAGTGG	AGGAGATGTT	2220
	GAGGCACGGC	TAGAAGGAGA	GCTGGTCCCC	TGTCCCCTGG	CAGAAGAGAA	CGAGTTCATC	2280
	CTGTACGCCA	CGCGCAAGTC	CATCCACCGC	TATGACCTGG	CTTCCGGAAC	CACGGAGCAG	2340
35	TTGCCCCTCA	CTGGGTTGCG	GGCAGCAGTG	GCCCTGGACT	TTGACTATGA	GCACAACTGC	2400
	CTGTATTGGT	CTGACCTGGC	CTTGGACGTC	ATCCAGCGCC	TCTGTTTGAA	CGGGAGTACA	2460
40	GGACAAGAGG	TGATCATCAA	CTCTGACCTG	GAGACGGTAG	AAGCTTTGGC	TTTTGAACCC	2520
	CTCAGCCAAT	TACTTTACTG	GGTGGACGCA	GGCTTTAAAA	AGATCGAGGT	AGCCAATCCA	2580
	GATGGTGACT	TCCGACTCAC	CGTCGTCAAT	TCCTCGGTGC	TGGATCGGCC	CCGGGCCCTG	2640
45	GTCCTTGTGC	CCCAAGAAGG	GATCATGTTC	TGGACCGACT	GGGGAGACCT	GAAGCCTGGG	2700
	ATTTATCGGA	GCAACATGGA	CGGATCTGCC	GCCTATCGCC	TCGTGTCGGA	GGATGTGAAG	276
	TGGCCCAATG	GCATTTCCGT	GGACGATCAG	TGGATCTACT	GGACGGATGC	CTACCTGGAC	2820
50	TGCATTGAGC	GCATCACGTT	CAGCGGCCAG	CAGCGCTCCG	TCATCCTGGA	CAGACTCCCG	2886
	CACCCCTATG	CCATTGCTGT	CTTTAAGAAT	GAGATTTACT	GGGATGACTG	GTCACAGCTC	294

	AGCATATTC	GAGCTTCTAA	GTACAGCGG	G TCCCAGATG(AGATTCTGGC	CAGCCAGCTC	3000
5	ACGGGGCTGA	TGGACATGAA	GATCTTCTAC	CAAGGGGAAGA	ACACAGGAAG	CAATGCGTGT	3060
	GTACCCAGGC	CGTGCAGCCT	GCTGTGCCT	G CCCAGAGCCA	ACAACAGCAA	AAGCTGCAGG	3120
	TGTCCAGATO	GCGTGGCCAG	CAGTGTCCTC	CCTTCCGGG	ACCTGATGTG	TGACTGCCCT	3180
10	AAGGGCTACO	AGCTGAAGAA	CAACACGTG1	GTCAAAGAAC	AAGACACCTG	TCTGCGCAAC	3240
	CAGTACCGCT	CCAGCAACGG	GAACTGCATC	CAACAGCATCT	GGTGGTGCGA	TTTCGACAAC	3300
	GACTGCGGAG	ACÁTGAGCGA	CGAGAAGAAC	TGCCCTACCA	CCATCTGCGA	CCTGGACACC	3360
15	CAGTTCCGTT	GCCAGGAGTC	TGGGACGTGC	ATCCCGCTCT	CCTACAAATG	TGACCTCGAG	3420
	GATGACTGTG	GGGACAACAG	TGACGAAAGG	CACTGTGAAA	TGCACCAGTG	CCGGAGCGAC	3480
20	GAATACAACT	GCAGCTCGGG	CATGTGCATC	CGCTCCTCCT	GGGTGTGCGA	CGGGGACAAC	3540
	GACTGCAGGG	ACTGGTCCGA	CGAGGCCAAC	TGCACAGCCA	TCTATCACAC	CTGTGAGGCC	3600
	TCCAACTTCC	AGTGCCGCAA	CGGGCACTGC	ATCCCCCAGC	GGTGGGCGTG	TGACGGCGAC	3660
25	GCCGACTGCC	AGGATGGCTC	TGATGAGGAT	CCAGCCAACT	GTGAGAAGAA	GTGCAACGGC	3720
	TTCCGCTGCC	CGAACGCCAC	CTGCATTCCC	TCCACCAAGC	ACTGTGACGG	CCTGCACGAT	3780
	TGCTCGGACG	GCTCCGACGA	GCAGCACTGC	GAGCCCCTGT	GTACACGGTT	CATGGACTTC	3840
30	GTGTGTAAGA	ACCGCCAGCA	GTGCCTCTTC	CACTCCATGG	TGTGCGATGG	GATCATCCAG	3900
	TGCCGTGACG	GCTCCGACGA	GGACCCAGCC	TTTGCAGGAT	GCTCCCGAGA	CCCCGAGTTC	3960
	CACAAGGTGT	GCGATGAGTT	CGGCTTCCAG	TGTCAGAACG	GCGTGTGCAT	CAGCTTGATC	4020
35	TGGAAGTGCG	ACGGGATGGA	TGACTGCGGG	GACTACTCCG	ACGAGGCCAA	CTGTGAAAAC	4080
	CCCACAGAAG	CCCCCAACTG	CTCCCGCTAC	TTCCAGTTCC	GGTGTGACAA	TGGCCACTGC	4140
10	ATCCCCAACA	GGTGGAAGTG	TGACAGGGAG	AATGACTGTG	GGGACTGGTC	CGACGAGAAG	4200
	GACTGTGGAG	ATTCACATGT	ACTTCCGTCT	ACGACTCCTG	CACCCTCCAC	GTGTCTGCCC	4260
	AATTACTACC	GCTGCGGCGG	GGGGGCCTGC	GTGATAGACA	CGTGGGTTTG	TGACGGGTAC	4320
5	CGAGATTGCG	CAGATGGATC	CGACGAGGAA	GCCTGCCCCT	CGCTCCCCAA	TGTCACTGCC	4380
	ACCTCCTCCC	CCTCCCAGCC	TGGACGATGC	GACCGATTTG	AGTTTGAGTG	CCACCAGCCA	4440
	AAGAAGTGCA	TCCCTAACTG	GAGACGCTGT	GACGGCCATC	AGGATTGCCA	GGATGGCCAG	4500
0	GACGAGGCCA	ACTGCCCCAC	TCACAGCACC	TTGACCTGCA	TGAGCTGGGA	GTTCAAGTGT	4560
	GAGGATGGCG	AGGCCTGCAT	CGTGCTGTCA	GAACGCTGCG	ACGGCTTCCT	GGACTGCTCA	4620

	GATGAGAGCG	ACGAGAAGGC	CTGCAGTGAT	GAGTTAACTG	TATACAAAGT	ACAGAATCTT	4680
	CAGTGGACAG	CTGACTTCTC	TGGGAATGTC	ACTTTGACCT	GGATGCGGCC	CAAAAAAATG	4740
	CCCTCTGCTG	CTTGTGTATA	CAACGTGTAC	TATAGAGTTG	TTGGAGAGAG	CATATGGAAG	4800
	ACTCTGGAGA	CTCACAGCAA	TAAGACAAAC	ACTGTATTAA	AAGTGTTGAA	ACCAGATACC -	4860
o	ACCTACCAGG	TTAAAGTGCA	GGTTCAGTGC	CTGAGCAAGG	TGCACAACAC	CAATGACTTT	4920
	GTGACCTTGA	GAACTCCAGA	GGGATTGCCA	GACGCCCCTC	AGAACCTCCA	GCTGTCGCTC	4980
	CACGGGGAAG	AGGAAGGTGT	GATTGTGGGC	CACTGGAGCC	CTCCCACCCA	CACCCACGGC	5040
5	CTCATTCGCG	AATACATTGT	AGAGTATAGC	AGGAGTGGTT	CCAAGGTGTG	GACTTCAGAA	5100
	AGGGCTGCTA	GTAACTTTAC	AGAAATAAAG	AACTTGTTGG	TCAACACCCT	GTACACCGTC	5160
	AGAGTGGCTG	CGGTGACGAG	TCGTGGGATA	GGAAACTGGA	GCGATTCCAA	ATCCATTACC	5220
20	ACCGTGAAAG	GAAAAGCGAT	CCCGCCACCA	AATATCCACA	TTGACAACTA	CGATGAAAAT	5280
	TCCCTGAGTT	TTACCCTGAC	CGTGGATGGG	AACATCAAGG	TGAATGGCTA	TGTGGTGAAC	5340
25	CTTTTCTGGG	CATTTGACAC	CCACAAACAA	GAGAAGAAAA	CCATGAACTT	CCAAGGGAGC	5400
	TCAGTGTCCC	ACAAAGTTGG	CAATCTGACA	GCACAGACGG	CCTATGAGAT	TTCCGCCTGG	5460
	GCCAAGACTG	ACTTGGGCGA	TAGTCCTCTG	TCATTTGAGC	ATGTCACGAC	CAGAGGGGTT	5520
30	CGCCCACCTG	CTCCTAGCCT	CAAGGCCAGG	GCTATCAATC	AGACTGCAGT	GGAATGCACC	5580
	TGGACAGGCC	CCAGGAATGT	GGTGTATGGC	ATTTTCTATG	CCACATCCTT	CCTGGACCTC	5640
	TACCGCAACC	CAAGCAGCCT	GACCACGCCG	CTGCACAACG	CAACCGTGCT	CGTCGGTAAG	5700
35	GATGAGCAGT	ATCTGTTTCT	GGTCCGGGTG	GTGATGCCCT	ACCAAGGGCC	GTCCTCGGAC	5760
	TACGTGGTCG	TGAAGATGAT	CCCGGACAGC	AGGCTTCCTC	CCCGCCACCT	GCATGCCGTT	5820
	CACACCGGCA	AGACCTCGGC	CGTCATCAAG	TGGGAGTCGC	CCTACGACTC	TCCTGACCAG	5880
4 0	GACCTGTTCT	ATGCGATCGC	AGTTAAAGAT	CTGATACGAA	AGACGGACCG	GAGCTACAAA	5940
	GTCAAGTCCC	GCAACAGCAC	CGTGGAGTAC	ACCCTGAGCA	AGCTGGAGCC	CGGAGGGAAA	6000
45	TACCACGTCA	TTGTGCAGCT	GGGGAACATG	AGCAAAGATG	CCAGTGTGAA	GATCACCACC	6060
	GTTTCGTTAT	CGGCACCCGA	TGCCTTAAAA	ATCATAACAG	AAAATGACCA	CGTCCTTCTC	6120
	TTCTGGAAAA	GTCTAGCTCT	AAAGGAAAAG	TATTTTAACG	AAAGCAGGGG	CTACGAGATA	6180
50	CACATGTTTG	ATAGCGCCAT	GAATATCACC	GCATACCTTG	GGAATACTAC	TGACAATTTC	6240
	TTTAAAATTT	CCAACCTGAA	GATGGGTCAC	AATTACACAT	TCACGGTCCA	GGCACGATGC	6300

	CTTT	rtgg	GCA (GCCA	GATC'	rg c	GGGG	AGCC	r GC	CGTG	CTAC	TGT	ATGA′	TGA	GCTG	GGGTCT	6360
5	GGTC	GCG	ATG (CGTC	GGCG	AT G	CAGG	CTGC	C AG	GTCT	ACTG	ATG'	rcgc	CGC	CGTG	GTGGTG	6420
•	CCCA	TCC	TGT '	TTCT	GATA	CT G	CTGA	GCCT	G GG	GGTC	GGGT	TTG	CCATO	CCT	GTAC	ACGAAG	6480
	CATO	CGGA	GGC '	TGCA	GAGC	AG C	TTCA	CCGC	C TT	CGCC	AACA	GCC	ACTAC	CAG	СТСС	AGACTC	6540
10	GGC1	rccg	CCA '	TCTT	CTCC	rc t	GGGG	ATGA(C TT	GGGG	GAGG	ATG	ATGA!	AGA	TGCT	CCTATG	6600
	ATCA	CTG	GAT 1	TTTC	GGAC	GA C	GTCC	CCATO	G GT	GATA	GCC						6639
	Seq	uen	ce 1	D N	o. :	2											
15	Len	gth	of	the	Sec	quer	ce:	221	L3								
	Тур	e: a	amir	no a	cid												
	Top	olo	gy:	lin	ear												
20	Mol	ecu.	lar	typ	e: 1	Prot	tein										
	Seq	uen	ce:														
25	Met	Ala	Thr	Arg	Ser	Ser	Arg	Arg	Glu	Ser	Arg	Leu	Pro	Phe	Leu	Phe	
					5					10					15		
	Thr	Leu	Val	Ala	Leu	Leu	Pro	Pro	Gly	Ala	Leu	Cys	Glu	Val	Trp	Thr	
30				20					25					30	1		
	Arg	Thr	Leu	His	Gly	Gly	Arg	Ala	Pro	Leu	Pro	Gln	Glu	Arg	Gly	Phe	
			35					40					45				
35	Arg	Val	Val	Gln	Gly	Asp	Pro	Arg	Glu	Leu	Arg	Leu	Trp	Glu	Arg	Gly	
		50					55					60					
10	Asp	Ala	Arg	Gly	Ala	Ser	Arg	Ala	Asp	Glu	Lys	Pro	Leu	Arg	Arg	Arg	
	65					70					75					80	
	Arg	Ser	Ala	Ala	Leu	Gln	Pro	Glu	Pro	Ile	Lys	Val	Tyr	Gly	Gln	Val	
1 5					85					90					95		
	Ser	Leu	Asn	Asp	Ser	His	Asn	Gln	Met	Val	Val	His	Trp	Ala	Gly	Glu	
				100					105					110			
50	Lys	Ser	Asn	Val	lle	Val	Ala	Leu	Ala	Arg	Asp	Ser	Leu	Ala	Leu	Ala	
			115					120					125				

	Arg	Pro	Arg	Ser	Ser	Asp	Val	Туг	Val	Ser	Туг	Asp	Tyr	Gly	Lys	Ser
5		130					135					140				
	Phe	Asn	Lys	He	Ser	Glu	Lys	Leu	Asn	Phe	Gly	Ala	Gly	Asn	Asn	Thr
	145					150					155					160
10	Glu	Ala	Val	Val	Ala	Gln	Phe	Туг	His	Ser	Pro	Ala	Asp	Asn	Lys	Arg
					165					170					175	
	Tyr	lle	Phe	Ala	Asp	Ala	Tyr	Ala	Gln	Tyr	Leu	Trp	lle	Thr	Phe	Asp
15				180					185					190		
	Phe	Cys	Asn	Thr	Ile	His	Gly	Phe	Ser	lle	Pro	Phe	Arg	Ala	Ala	Asp
			195					200					205			
20	Leu	Leu	Leu	His	Ser	Lys	Ala	Ser	Asn	Leu	Leu	Leu	Gly	Phe	Asp	Arg
		210					215					220				
25	Ser	His	Pro	Asn	Lys	Gln	Leu	Trp	Lys	Ser	Asp	Asp	Phe	Gly	Gln	Thr
	225					230					235					240
	Trp	lle	Met	Ile	Gln	Glu	His	Val	Lys	Ser	Phe	Ser	Trp	Gly	Ile	Asp
30					245					250					255	
	Pro	Tyr	Asp	Lys	Pro	Asn	Thr	Ile	Туг	Ile	Glu	Arg	His	Glu	Pro	Ser
•				260					265					270		
35	Gly	Tyr	Ser	Thr	Val	Phe	Arg	Ser	Thr	Asp	Phe	Phe	Gl n	Ser	Arg	Glu
			275					280					285			
4 0	Asn	Gln	Glu	Val	He	Leu	Glu	Glu	Val	Arg	Asp	Phe	Gln	Leu	Arg	Asp
		290					295					300				
,	Lys	Tyr	Met	Phe	Ala	Thr	Lys	Val	Val	His	Leu	Leu	Gly	Ser	Pro	Leu
45	305					310					315					320
	Gln	Ser	Ser	Val	Gln	Leu	Trp	Val	Ser	Phe	Gly	Arg	Lys	Pro	Met	Arg
					325					330					335	
50	Ala	Ala	Gln	Phe	Val	Thr	Arg	His	Pro	He	Asn	Glu	Туг	Tyr	lle	Ala
				340					345					350		

	Asp	Ala	Ser	Glu	Asp	Gln	Val	Phe	Val	Cys	Val	Ser	His	Ser	Asn	Asn
5			355					360					365			
5	Arg	Thr	Asn	Leu	Tyr	lle	Ser	Glu	Ala	Glu	Gly	Leu	Lys	Phe	Ser	Leu
		370					375					380				
10	Ser	Leu	Glu	Asn	Val	Leu	Туг	Tyr	Thr	Pro	Gly	Gly	Ala	Gly	Ser	Asp
	385					390					395					400
	Thr	Leu	Val	Arg	Tyr	Phe	Ala	Asn	Glu	Pro	Phe	Ala	Asp	Phe	His	Arg
15					405					410	•				415	
	Val	Glu	Gly	Leu	Gln	Gly	Val	Туг	Ιlе	Ala	Thr	Leu	He	Asn	Gly	Ser
				420					425					430		
20	Met	Asn	Glu	Glu	Asn	Met	Arg	Ser	Val	Ile	Thr	Phe	Asp	Lys	Gly	Gly
			435					440					445			
25	Thr	Trp	Glu	Phe	Leu	Gln	Ala	Pro	Ala	Phe	Thr	Gly	Туг	Gly	Glu	Lys
		450					455					460				
	He	Asn	Cys	Glu	Leu	Ser	Glu	Gly	Cys	Ser	Leu	His	Leu	Ala	Gln	Arg
30	465					470					475					480
	Leu	Ser	Gln	Leu	Leu	Asn	Leu	Gln	Leu	Arg	Arg	Met	Pro	Ile	Leu	Ser
					485					490					495	
35	Lys	Glu	Ser	Ala	Pro	Gly	Leu	Пe	Ile	Ala	Thr	Gly	Ser	Val	Gly	Lys
				500					505					510		
	Asn	Leu	Ala	Ser	Lys	Thr	Asn	Val	Туг	lle	Ser	Ser	Ser	Ala	Gly	Ala
40			515					520					525			
	Arg	Trp	Arg	Glu	Ala	Leu	Pro	Gly	Pro	His	Туг	Tyr	Thr	Trp	Gly	Asp
45		530					535					540				
	His	Gly	Gly	lle	He	Met	Ala	Ile	Ala	Gln	Gly	Met	Glu	Thr	Asn	Glu
	545					550					555					560
50	Leu	Lys	Туг	Ser	Thr	Asn	Glu	Gly	Glu	Thr	Trp	Lys	Ala	Phe	Thr	Phe
					565					570					575	

	Ser	Glu	Lys	Pro	Val	Phe	Val	Туr	Gly	Leu	Leu	Thr	Glu	Pro	Gly	Glu
5				580					585					590		
	Lys	Ser	Thr	Val	Phe	Thr	Ιlе	Phe	Gly	Ser	Asn	Lys	Glu	Asn	Val	His
			595					600					605			
10	Ser	Trp	Leu	He	Leu	Gln	Val	Asn	Ala	Thr	Asp	Ala	Leu	Gly	Val	Pro
		610					615					620				
	Cys	Thr	Glu	Asn	Asp	Туг	Lys	Leu	Trp	Ser	Pro	Ser	Asp	Glu	Arg	Gly
15 .	625					630					635					640
	Asn	Glu	Cys	Leu	Leu	Gly	His	Lys	Thr	Val	Phe	Lys	Arg	Arg	Thr	Pro
20					645		-			650					655	
	His	Ala	Thr	Cys	Phe	Asn	Gly	Glu	Asp	Phe	Asp	Arg	Pro	Val	Val	Val
				660					665					670		
25	Ser	Asn	Cys	Ser	Cys	Thr	Arg	Glu	Asp	Tyr	Glu	Cys	Asp	Phe	Gly	Phe
			675					680					685			
	Arg	Met	Ser	Glu	Asp	Leu	Ala	Leu	Glu	Val	Cys	Val	Pro	Asp	Pro	Gly
30		690					695					700				
	Phe	Ser	Gly	Lys	Ser	Ser	Pro	Pro	Val	Pro	Cys	Pro	Val	Gly	Ser	Thr
35	705					710				•	715					720
	Туг	Arg	Arg	Ser	Arg	Gly	Tyr	Arg	Lys	Ile	Ser	Gly	Asp	Thr	Cys	Ser
					725					730					735	
40	Gly	Gly	Asp	Val	Glu	Ala	Arg	Leu	Glu	Gly	Glu	Leu	Val	Pro	Cys	Pro
				740					745					750		
	Leu	Ala	Glu	Glu	Asn	Glu	Phe	lle	Leu	Tyr	Ala	Thr	Arg	Lys	Ser	Пе
4 5			755					760					765			
	His	Arg	Tyr	Asp	Leu	Ala	Ser	Gly	Thr	Thr	Glu	Gln	Leu	Pro	Leu	Thr
		770					775					780				
50	Gly	Leu	Arg	Ala	Ala	Val	Ala	Leu	Asp	Phe	Asp	Туг	Glu	His	Asn	Cys
	785					790					795					800

	Leu	Tyr	Trp	Ser	Asp	Leu	Ala	Leu	Asp	Val	He	Gln	Arg	Leu	Cys	Leu
•					805					810					815	
	Asn	Gly	Ser	Thr	Gly	Gln	Glu	Val	Ile	lle	Asn	Ser	Asp	Leu	Glu	Thr
				820					825					830		
10	Val	Glu	Ala	Leu	Ala	Phe	Glu	Pro	Leu	Ser	Gln	Leu	Leu	Туг	Trp	Val
			835					840					845			
	Asp	Ala	Gly	Phe	Lys	Lys	lle	Glu	Val	Ala	Asn	Pro	Asp	Gly	Asp	Phe
15		850					855					860				
	Arg	Leu	Thr	Val	Val	Asn	Ser	Ser	Val	Leu	Asp	Arg	Pro	Arg	Ala	Leu
20	865					870					875					880
	Val	Leu	Val	Pro	Gln	Glu	Gly	He	Met	Phe	Тгр	Thr	Asp	Trp	Gly	Asp
					885					890					895	
25	Leu	Lys	Pro	Gly	Ile	Tyr	Arg	Ser	Asn	Met	Asp	Gly	Ser	Ala	Ala	Туг
				900					905					910		
	Arg	Leu	Val	Ser	Glu	Asp	Val	Lys	Trp	Pro	Asn	Gly	He	Ser	Val	Asp
30			915					920					925	i		
	Asp	Gln	Trp	Ιle	Туг	Trp	Thr	Asp	Ala	Туг	Leu	Asp	Cys	lle	Glu	Arg
35		930					935					940				
	lle	Thr	Phe	Ser	Gly	Gln	Gln	Arg	Ser	Val	He	Leu	Asp	Arg	Leu	Pro
	945					950					955					960
40	His	Pro	Tyr	Ala	Ile	Ala	Val	Phe	Lys	Asn	Glu	He	Туг	Trp	Asp	Asp
•					965	j				970					975	
	Trp	Ser	Gln	Leu	Ser	Ile	Phe	Arg	Ala	Ser	Lys	Туг	Ser	Gly	' Ser	Gln
45				980)				985	j				990	j	
	Met	Glu	Ile	Leu	Ala	Ser	Gln	Leu	Thr	Gly	Leu	Met	. Asp	Met	. Lys	lle
			995	5				1000)				1005	5		
50	Phe	Tyr	Lys	Gly	Lys	s Asr	Thr	Gly	Ser	Asn	Ala	Cys	s Val	l Pro) Arg	Pro
		1010)				1015	5				1020)			

	Cy:	s Se	r Le	u Le	u Cy	s Le	u Pr	o Ar	g Al	a As	n Asr	Sei	r Ly	s Se	r Cy	s Arg
5	10	25				10	30				103	35				1040
·	Cys	s Pr	o As	p Gl	y Va	1 A]	a Se	r Se	r Va	l Lei	u Pro	Ser	r Gl	y As	p Le	u Met
						45				109					10	
10	Cys	s Ası	р Су	s Pr	o Ly	s Gl	у Ту	r Gl	u Lei	ı Lys	s Asn	Asr	1 Th	r Cy:	s Va	l Lys
				10					106					10		
	Glu	ı Glu	ı As	p Th	r Cy:	s Lei	ı Arı	g Ası	n Glr	туг	Arg	Cys	Sei	r Ası	n Gly	y Asn
15			10					108					108			
	Cys	He	e Ası	n Sei	- 11	e Trp	Trp	Cys	s Asp	Phe	e Asp	Asn	. Ası	Cys	s Gly	/ Asp
		109					109					110				•
20	Met	Ser	Ası	Glu	ı Lys	s Asr	Cys	Pro	Thr	Thr	lle	Cys	Ası	Lei	ı Asp	Thr
	110					111					111					1120
25	Gln	Phe	: Arg	Cys	Glr	ı Glu	Ser	Gly	Thr	Cys	Ile	Pro	Leu	ı Ser	Туг	Lys
					112					113					113	
	Cys	Asp	Leu	Glu	Asp	Asp	Cys	Gly	Asp	Asn	Ser	Asp	Glu	Arg		
30				114					114					115		•
•	Glu	Met	His	Gln	Cys	Arg	Ser	Asp	Glu	Tyr	Asn	Cys	Ser			Met
			115					116					116		·	
35	Cys	He	Arg	Ser	Ser	Trp	Val	Cys	Asp	Gly	Asp	Asn	Asp	Cys	Arg	Asp
		1170					1175					1180				-
40	Trp	Ser	Asp	Glu	Ala	Asn	Cys	Thr	Ala	lle	Tyr	His	Thr	Cys	Glu	Ala
40	1185					1190					1195					1200
	Ser	Asn	Phe	Gln	Cys	Arg	Asn	Gly	His	Cys	Ile I	Pro	Gln	Arg	Тгр	
45					1205					1210					1215	
	Cys	Asp	Gly	Asp	Ala	Asp	Cys	Gln	Asp	Gly	Ser A	Asp (Glu	Asp		
				1220					1225					1230		
50	Asn	Cys	Glu	Lys	Lys	Cys	Asn	Gly	Phe	Arg	Cys f	Pro /	Asn	Gly	Thr	Cys
			1235					1240					1245			-

	Ile	Pro	Ser	Thr	Lys	His	Cys	Asp	Gly	Leu	His	Asp	Cys	Ser	Asp	Gly
		1250)				1255	j				1260)			
5	Ser	Asp	Glu	Gln	His	Cys	Glu	Pro	Leu	Cys	Thr	Arg	Phe	Met	Asp	Phe
	1265	5				1270)				1275	5				1280
	Val	Cys	Lys	Asn	Arg	Gln	Gln	Cys	Leu	Phe	His	Ser	Met	Val	Cys	Asp
10					1285	i				1290)				1295	i
	Gly	lle	Ile	Gln	Cys	Arg	Asp	Gly	Ser	Asp	Glu	Asp	Pro	Ala	Phe	Ala
15				1300)				1305	5				1310)	
	Gly	Cys	Ser	Arg	Asp	Pro	Glu	Phe	His	Lys	Val	Cys	Asp	Glu	Phe	Gly
			1315	5				1320)				1325	5		
20	Phe	Gln	Cys	Gln	Asn	Gly	Val	Cys	lle	Ser	Leu	lle	Trp	Lys	Cys	Asp
		1330)				1335	5				1340)			
	Gly	Met	Asp	Asp	Cys	Gly	Asp	Tyr	Ser	Asp	Glu	Ala	Asn	Cys	Glu	Asn
25	1345	5				1350)				135	5				1360
	Pro	Thr	Glu	Ala	Pro	Asn	Cys	Ser	Arg	Tyr	Phe	Gln	Phe	Arg	Cys	Asp
					1365	5				1370)				1375	5
30	Asn	Gly	His	Cys	Ile	Pro	Asn	Arg	Trp	Lys	Cyș	Asp	Arg	Glu	Asn	Asp
				138)				138	5				139	0	
35	Cys	Gly	Asp	Trp	Ser	Asp	Glu	Lys	Asp	Cys	Gly	Asp	Ser	His	Val	Leu
			139	5				140	0				140	5		
	Pro	Ser	Thr	Thr	Pro	Ala	Pro	Ser	Thr	Cys	Leu	Pro	Asn	Tyr	Tyr	Arg
40		141	0				141	5				142	0			
	Cys	Gly	Gly	Gly	Ala	Cys	Val	lle	Asp	Thr	Trp	Val	Cys	Asp	Gly	Tyr
	142	5				143	0				143	5				1440
45	Arg	Asp	Cys	Ala	Asp	Gly	Ser	Asp	Glu	Glu	Ala	Cys	Pro	Ser	Leu	Pro
					144	5				145	0				145	5
	Asn	Val	Thr	Ala	Thr	Ser	Ser	Pro	Ser	Gln	Pro	Gly	Arg	Cys	Asp	Arg
50				146	0				146	55				147	0	

	Phe	Glu	Phe	Glu	Cys	His	Gln	Pro	Lys	Lys	Cys	He	Pro	Asn	Trp	Arg
5			147	5				148	0				148	5		
	Arg	Cys	Asp	Gly	His	Gln	Asp	Cys	Gln	Asp	Gly	Gln	Asp	Glu	Ala	Asn
		149	0				149	5				150	0			
10	Cys	Pro	Thr	His	Ser	Thr	Leu	Thr	Cys	Met	Ser	Trp	Glu	Phe	Lys	Cys
	150	5				151	0				151	5				1520
	Glu	Asp	Gly	Glu	Ala	Cys	lle	Val	Leu	Ser	Glu	Arg	Cys	Asp	Gly	Phe
15					152	5				153)				1539	ō
	Leu	Asp	Cys	Ser	Asp	Glu	Ser	Asp	Glu	Lys	Ala	Cys	Ser	Asp	Glu	Leu
20				154	0				154	5				155	0	
	Thr	Val	Туг	Lys	Val	Gln	Asn	Leu	Gln	Trp	Thr	Ala	Asp	Phe	Ser	Gly
			155	5				1560)				156	5		
25	Asn	Va!	Thr	Leu	Thr	Trp	Met	Arg	Pro	Lys	Lys	Met	Pro	Ser	Ala	Ala
		1570)			•	157	5				158	0			
	Cys	Va!	Туг	Asn	Val	Туг	Tyr	Arg	Val	Val	Gly	Glu	Ser	Пe	Trp	Lys
30	1585	5				1590)				1595	5				1600
· .	Thr	Leu	Glu	Thr	His	Ser	Asn	Lys	Thr	Asn	Thr	Val	Leu	Lys	Val	Leu
35					1605	,				1610)				1615	5
	Lys	Pro	Asp	Thr	Thr	Туг	Gln	Val	Lys	Val	Gln	Val	Gln	Cys	Leu	Ser
				1620	١				1625	5				1630)	
40	Lys	Val			Thr	Asn	Asp	Phe	Val	Thr	Leu	Arg	Thr	Pro	Glu	Gly
			1635					1640					1645			
	Leu			Ala	Pro	Gln			Gln	Leu	Ser	Leu	His	Gly	Glu	Glu
45		1650					1655					1660				
			Val	He				Тгр	Ser	Pro	Pro	Thr	His	Thr	His	Gly
50	1665					1670					1675					1680
	Leu	He	Arg	Glu			Val	Glu	Tyr	Ser	Arg	Ser	Gly	Ser	Lys	Val
					1685					1690					1695	

	Trp	Thr	Ser	Glu	Arg	Ala	Ala	Ser	Asn	Phe	Thr	Glu	lle	Lys	Asn	Leu
5				.1700)				1705	5				1710)	
	Leu	Val	Asn	Thr	Leu	Tyr	Thr	Val	Arg	Val	Ala	Ala	Val	Thr	Ser	Arg
			1715	õ				1720)				1725	5		
10	Gly	lle	Gly	Asn	Trp	Ser	Asp	Ser	Lys	Ser	lle	Thr	Thr	Val	Lys	Gly
		1730)				1735	5				1740)			
	Lys	Ala	He	Pro	Pro	Pro	Asn	Ile	His	Ile	Asp	Asn	Tyr	Asp	Glu	Asn
15	1745	5				1750)				175	5				1760
	Ser	Leu	Ser	Phe	Thr	Leu	Thr	Val	Asp	Gly	Asn	Ile	Lys	Val	Asn	Gly
00					176	5				1770)				177	5
20	Tyr	Val	Val	Asn	Leu	Phe	Trp	Ala	Phe	Asp	Thr	His	Lys	Gln	Glu	Lys
				178	0				178	5	•			179	0	
25	Lys	Thr	Met	Asn	Phe	Gln	Gly	Ser	Ser	Val	Ser	His	Lys	Val	Gly	Asn
			179	5				180	0				180	5		
	Leu	Thr	Ala	Gln	Thr	Ala	Туг	Glu	Ile	Ser	Ala	Trp	Ala	Lys	Thr	Asp
30		181	0				181	5				182	0			
	Leu	Gly	Asp	Ser	Pro	Leu	Ser	Phe	Glu	His	Val	Thr	Thr	Arg	Gly	Val
	1825	5				1830)				1835	5				1840
35	Arg	Pro	Pro	Ala	Pro	Ser	Leu	Lys	Ala	Arg	Ala	Ile	Asn	Gln	Thr	Ala
					1845	5				1850)				1855	5
40	Val	Glu	Cys	Thr	Trp	Thr	Gly	Pro	Arg	Asn	Val	Val	Tyr	Gly	He	Phe
				1860)				1869	5				1870)	
	Tyr	Ala	Thr	Ser	Phe	Leu	Asp	Leu	Туг	Arg	Asn	Pro	Ser	Ser	Leu	Thr
45			1875	5				1880)				1885	5		
	Thr	Pro	Leu	His	Asn	Ala	Thr	Val	Leu	Val	Gly	Lys	Asp	Glu	Gln	Tyr
		1890	0				189	5				1900	0			
50	Leu	Phe	Leu	Val	Arg	Val	Val	Met	Pro	Tyr	Gln	Gly	Pro	Ѕег	Ser	Asp
	1905	5				1910	0				191	5				1920

	Tyr	Val	Val	Val	Lys	Met	Ile	Pro) Asp	Ser	Arg	Leu	Pro	Pro	Arg	His
5					192	5				193	0				193	5
	Leu	His	Ala	Val	His	Thr	Gly	Lys	Thr	Ser	Ala	Val	Ιlе	Lys	Тгр	Glu
				194	0				194	5				195	0	
10	Ser	Pro	Tyr	Asp	Ser	Pro	Asp	Gln	Asp	Leu	Phe	Туг	Ala	Ile	Ala	Val
			195	5				196	0				196	5		
	Lys	Asp	Leu	Ιle	Arg	Lys	Thr	Asp	Arg	Ser	Туг	Lys	Val	Lys	Ser	Arg
15		197	0				197	5				198	0			
	Asn	Ser	Thr	Val	Glu	Tyr	Thr	Leu	Ser	Lys	Leu	Glu	Pro	Gly	Gly	Lys
20	198	5				199	0				199	5				2000
	Tyr	His	Val	Ile	Val	Gln	Leu	Gly	Asn	Met	Ser	Lys	Asp	Ala	Ser	Val
					2009	5				201	0			•	2019	õ
25	Lys	He	Thr	Thr	Val	Ser	Leu	Ser	Ala	Pro	Asp	Ala	Leu	Lys	Ile	Ιle
				2020	0				202	5				2030)	
	Thr	Glu	Asn	Asp	His	Val	Leu	Leu	Phe	Trp	Lys	Ser	Leu	Ala	Leu	Lys
30			203	5				204	0				204	5		
	Glu	Lys	Tyr	Phe	Asn	Glu	Ser	Arg	Gly	Tyr	Glu	lle	His	Met	Phe	Asp
o E		2050)				2055	5				2060)			
25	Ser	Ala	Met	Asn	Пе	Thr	Ala	Tyr	Leu	Gly	Asn	Thr	Thr	Asp	Asn	Phe
	2065	5				2070)				2075	5				2080
0	Phe	Lys	lle	Ser	Asn	Leu	Lys	Met	Gly	His	Asn	Tyr	Thr	Phe	Thr	Val
					2085	•				2090)				2095	5
	Gln	Ala	Arg	Cys	Leu	Leu	Gly	Ser	Gln	Пе	Cys	Gly	Glu	Pro	Ala	Val
5				2100)				2105	5				2110	1	
	Leu	Leu	Tyr	Asp	Glu	Leu	Gly	Ser	Gly	Gly	Asp	Ala	Ser	Ala	Met	Gln
			2115	5				2120)				2125	i		
o	Ala	Ala	Arg	Ser	Thr	Asp	Val	Ala	Ala	Val	Val	Val	Pro	lle	Leu	Phe
		2130)				2135	i				2140)			

	Leu lle Leu Leu	Ser Leu Gly	Val Gly	Phe Ala	lle Leu T	yr Thr Lys
5	2145	2150		215	5	2160
	His Arg Arg Leu	Gln Ser Ser	Phe Thr	Ala Phe	Ala Asn S	er His Tyr
		2165		2170		2175
10	Ser Ser Arg Leu	Gly Ser Ala	lle Phe	Ser Ser	Gly Asp A	sp Leu Gly
	2180	1	2185	5	2	190
	Glu Asp Asp Glu	Asp Ala Pro	Met Ile	Thr Gly	Phe Ser A	sp Asp Val
15	2195		2200		2205	
	Pro Met Val Ile	Ala				
20	2210					
20	Sequence ID No					
	Length of the		6961			
25	Type: nucleic					
	Strandedness:					
	Topology: line	ar				
30	Molecular type	: cDNA to	mRNA			
	Feature:					
	Name/Key: sig					
35	Location: 178					
	Identification		S			
40	Name/Key: mat	_ ~				
	Location: 262	6816				
	Identification	n method:	S			
45	Sequence:					
	CCGCGAGCCG CACACG	TGAC GGCGCCG	CGC CGCG	CCGCGC C	GCGCCGAGC	GGGACCCAGC 60
	GGCTGCCCGG AGCCCC	GGGA GCGGCGC	GCG CGCG	GCCCCG G	CCCCGCCGC	TCGGCCGGCG 120
50	GCGCGCTGCA CATTCT	CTCC TGGCGGC	GGC GCCA	CCTGCA G	CCGCGTTCG	CCCGAACATG 180
						Met

																	i	
_	GCG	ACA	CGG	AGC	AGC	AGG	AGG	GAG	TCG	CGA	СТС	CCC	TTC	СТА	TTC	ACC		228
5	Ala	Thr	Arg	Ser	Ser	Arg	Arg	Glu	Ser	Arg	Leu	Pro	Phe	Leu	Phe	Thr		
				5					10					15				
10	CTG	GTC	GCG	CTG	CTG	CCG	CCC	GGG	GCT	CTC	TGC	GAG	GTG	TGG	ACG	CGG	-	276
	Leu	Val	Ala	Leu	Leu	Pro	Pro	Gly	Ala	Leu	Cys	Glu	Val	Trp	Thr	Arg		
			20					25					30					
15	ACA .	CTG	CAC	GGC	GGC	CGC	GCG	CCC	TTA	CCC	CAG	GAG	CGG	GGC	TTC	CGC		324
	Thr	Leu	His	Gly	Gly	Arg	Ala	Pro	Leu	Pro	Gln	Glu	Arg	Gly	Phe	Arg		
		35					4 0					45						
20	GTG	GTG	CAG	GGC	GAC	CCG	CGC	GAG	CTG	CGG	CTG	TGG	GAG	CGC	GGG	GAT		372
	Val	Val	Gln	Gly	Asp	Pro	Arg	Glu	Leu	Arg	Leu	Trp	Glu	Arg	Gly	Asp		
25	50					55			_		60					65		
	GCC	AGG	GGG	GCG	AGC	CGG	GCG	GAC	GAG	AAG	CCG	СТС	CGG	AGG	AGA	CGG		420
	Ala	Arg	Gly	Ala	Ser	Arg	Ala	Asp	Glu	Lys	Pro	Leu	Arg	Arg	Arg	Arg		
30					70					7 5					80			
	AGC	GCT	GCC	CTG	CAG	CCC	GAG	CCC	ATC	AAG	GTG	TAC	GGA	CAG	GTC	AGC		468
	Ser	Ala	Ala	Leu	Gln	Pro	Glu	Pro	lle	Lys	Val	Туг	Gly	Gln	Val	Ser		
35				85					90					95				
	CTC	AAT	GAT	TCC	CAC	AAT	CAG	ATG	GTG	GTG	CAC	TGG	GCC	GGA	GAG	AAA		516
40	Leu	Asn	Asp	Ser	His	Asn	Gln	Met	Val	Val	His	Trp	Ala	Gly	Glu	Lys		
			100					105					110					
	AGC	AAC	GTG	ATC	GTG	GCC	TTG	GCC	CGG	GAC	AGC	CTG	GCG	TTG	GCC	AGG		564
45	Ser	Asn	Val	·Ile	Val	Ala	Leu	Ala	Arg	Asp	Ser	Leu	Alà	Leu	Ala	Arg		
		115					120					125						
	CCC	AGG	AGC	AGT	GAT	GTG	TAC	GTG	TCT	TAT	GAC	TAT	GGA	AAA	TCA	TTC		612
50	Pro	Arg	Ser	Ser	Asp	Val	Туг	Val	Ser	Туг	Asp	Tyr	Gly	Lys	Ser	Phe		
	130					135					140					145		

	AAT	AAG	ATT	TCA	GAG	AAA	TTG	AAC	TTC	GGC	GCG	GGA	AAT	AAC	ACA	GAG	660
5	Asn	Lys	He	Ser	Glu	Lys	Leu	Asn	Phe	Gly	Ala	Gly	Asn	Asn	Thr	Glu	
					150					155					160		
	GCT	GTG	GTG	GCC	CAG	TTC	TAC	CAC	AGC	CCT	GCG	GAC	AAC	AAA	CGG	TAC	708
10	Ala	Val	Val	Ala	Gln	Phe	Туг	His	Ser	Pro	Ala	Asp	Asn	Lys	Arg	Tyr	
				165					170					175			
	ATC	TTC	GCA	GAT	GCC	TAC	GCC	CAG	TAT	CTC	TGG	ATC	ACG	TTT	GAC	TTC	756
15	Ile	Phe	Ala	Asp	Ala	Tyr	Ala	Gln	Tyr	Leu	Trp	He	Thr	Phe	Asp	Phe	
			180					185					190				
	TGC	AAC	ACC	ATC	CAT	GGC	TTT	TCC	ATC	CCG	TTC	CGG	GCA	GCT	GAT	СТС	804
20	Cys	Asn	Thr	lle	His	Gly	Phe	Ser	Ile	Pro	Phe	Arg	Ala	Ala	Asp	Leu	
		195					200					205					
25	СТА	CTC	CAC	AGT	AAG	GCC	TCC	AAC	CTT	СТС	CTG	GGC	TTC	GAC	AGG	TCT	852
	Leu	Leu	His	Ser	Lys	Ala	Ser	Asn	Leu	Leu	Leu	Gly	Phe	Asp	Arg	Ser	
	210					215					2 <u>2</u> 0					225	
30	CAC	CCC	AAC	AAG	CAG	CTG	TGG	AAG	TCG	GAT	GAT	TTT	GGC	CAG	ACC	TGG	900
	His	Pro	Asn	Lys	Gln	Leu	Trp	Lys	Ser	Asp	Asp	Phe	Gly	Gln	Thr	Trp	
					230					235					240		
35	ATC	ATG	ATT	CAA	GAA	CAC	GTG	AAG	TCC	TTT	TCT	TGG	GGA	ATT	GAT	CCC	948
	Ile	Met	He	Gln	Glu	His	Val	Lys	Ser	Phe	Ser	Тгр	Gly	He	Asp	Pro	
				245					250					255			
40	ТАТ	GAC	AAA	CCA	AAC	ACC	ATC	TAC	ATC	GAA	CGG	CAC	GAA	ССТ	тст	GGC	996
	Туг	Asp	Lys	Pro	Asn	Thr	He	Туг	Ιle	Glu	Arg	His	Glu	Pro	Ser	Gly	
45	`		260					265					270				
	TAC	TCC	ACG	GTT	TTC	CGA	AGT	ACA	GAC	TTC	TTC	CAG	TCC	CGG	GAA	AAC	1044
	Tyr	Ser	Thr	Val	Phe	Arg	Ser	Thr	Asp	Phe	Phe	Gln	Ser	Arg	Glu	Asn	
50		275					280					285					
	CAG	GAA	GTG	ATC	TTG	GAG	GAA	GTG	AGA	GAC	TTT	CAG	CTT	CGG	GAC	AAG	1092

	GIn	Glu	Val	He	Leu	Glu	Glu	Val	Arg	Asp	Phe	Gln	Leu	Arg	Asp	Lys	
5	290					295					300					305	
	TAC	ATG	TTT	GCT	ACA	AAG	GTG	GTG	CAT	CTC	TTG	GGC	AGT	CCA	CTG	CAG	1140
	Tyr	Met	Phe	Ala	Thr	Lys	Val	Val	His	Leu	Leu	Gly	Ser	Pro	Leu	Gln	
10					310					315					320		
	TCT	TCT	GTC	CAG	CTC	TGG	GTC	TCC	TTT	GGC	CGG	AAG	CCC	ATG	CGG	GCC	1188
	Ser	Ser	Val	Gln	Leu	Trp	Val	Ser	Phe	Gly	Arg	Lys	Pro	Met	Arg	Ala	
15				325					330					335			
	GCC	CAG	TTT	GTT	ACA	AGA	CAT	CCT	ATC	AAC	GAA	TAT	TAC	ATC	GCG	GAT	1236
00	Ala	Gln	Phe	Val	Thr	Arg	His	Pro	Ile	Asn	Glu	Tyr	Tyr	lle	Ala	Asp	
20			340					345					350				
	GCC	TCG	GAG	GAC	CAG	GTG	TTT	GTG	TGT	GTC	AGT	CAC	AGC	AAC	AAC	CGC	1284
25	Ala	Ser	Glu	Asp	Gln	Val	Phe	Val	Cys	Val	Ser	His	Ser	Asn	Asn	Arg	
		355					360					365					
	ACC	AAC	СТС	TAC	ATC	TCG	GAG	GCA	GAG	GGC	TTG	AAG	TTC	тст	CTG	TCC	1332
30	Thr	Asn	Leu	Туг	Пе	Ser	Glu	Ala	Glu	Gly	Leu	Lys	Phe	Ser	Leu	Ser	
	370					375					380					385	
	CTG	GAG	AAC	GTG	CTC	TAC	TAC	ACC	CCG	GGA	GGG	GCC	GGC	AGT	GAC	ACC	1380
35	Leu	Glu	Asn	Val	Leu	Tyr	Туг	Thr	Pro	Gly	Gly	Ala	Gly	Ser	Asp	Thr	
•			-		390					395					400		
40	TTG	GTG	AGG	TAC	TTT	GCA	AAT	GAA	CCG	TTT	GCT	GAC	TTC	CAT	CGT	GTG	1428
40	Leu	Val	Arg	Туг	Phe	Ala	Asn	Glu	Pro	Phe	Ala	Asp	Phe	His	Arg	Val	
				405					410					415			
4 5	GAA	GGG	TTG	CAG	GGA	GTC	TAC	АТТ	GCT	ACT	CTG	ATT	AAT	GGT	тст	ATG	1476
	Glu	Gly	Leu	Gln	Gly	Val	Tyr	He	Ala	Thr	Leu	Ile	Asn	Gly	Ser	Met	
			420					425					430				
50	AAT	GAG	GAG	AAC	ATG	AGA	тст	GTC	ATC	ACC	TTT	GAC	AAA	GGG	GGC	ACC	1524
	Asn	Glu	Glu	Asn	Met	Arg	Ser	Val	lle	Thr	Phe	Asp	Lys	Gly	Gly	Thr	

		435					440				4	145			•		
5	TGG	GAA	TTT	CTG	CAG	GCT	CCA	GCC	TTC	ACG	GGG	TAT	GGA	GAG	AAA	ATC	1572
	Trp	Glu	Phe	Leu	Gln	Ala	Pro	Ala	Phe	Thr	Gly	Туг	Gly	Glu	Lys	lle	
	450					455					460					465	
10	AAC	TGT	GAG	CTG	TCC	GAG	GGC	TGT	TCC	CTC	CAC	CTG	GCC	CAG	CGC	CTC	1620
	Asn	Cys	Glu	Leu	Ser	Glu	Gly	Cys	Ser	Leu	His	Leu	Ala	Gln	Arg	Leu	
					470					475					480		
15	AGC	CAG	CTG	CTC	AAC	СТС	CAG	СТС	CGG	AGG	ATG	CCC	ATC	CTG	TCC	AAG	1668
	Ser	Gln	Leu	Leu	Asn	Leu	Gln	Leu	Arg	Arg	Met	Pro	lle	Leu	Ser	Lys	
20				485					490					495			
20	GAG	TCG	GCG	CCT	GGC	CTC	ATC	ATT	GCC	ACG	GGC	TCA	GTG	GGA	AAG	AAC	1716
	Glu	Ser	Ala	Pro	Gly	Leu	Ile	Ιlе	Ala	Thr	Gly	Ser	Val	Gly	Lys	Asn	
25			500					505					510				
	TTG	GCT	AGC	AAG	ACA	AAC	GTG	TAC	ATC	TCT	AGC	AGT	GCT	GGA	GCC	AGG	1764
	Leu	Ala	Ser	Lys	Thr	Asn	Val	Туг	Ιle	Ser	Ser	Ser	Ala	Gly	Ala	Arg	-
30		515					520					525					
	TGG	CGA	GAG	GCA	CTT	CCT	GGA	CCT	CAC	TAC	TÀT	ACA	TGG	GGA	GAC	CAT	1812
	Trp	Arg	Glu	Ala	Leu	Pro	Gly	Pro	His	Туг	Туг	Thr	Trp	Gly	Asp	His	
35	530					535					540					545	
	GGC	GGC	ATC	ATC	ATG	GCC	ATT	GCC	CAA	GGC	ATG	GAA	ACC	AAC	GAA	CTG	1860
40	Gly	Gly	lle	Ile	Met	Ala	Ile	Ala	Gln	Gly	Met	Glu	Thr	Asn	Glu	Leu	
					550					555	5				560	1	
	AAG	TAC	AGT	ACC	AAC	GAA	GGG	GAG	ACC	TGC	AAA	GCC	TTC	ACC	T TC	TCT	1908
45	Lys	Tyr	Ser	Thr	Asn	Glu	Gly	Glu	Thr	Trp	Lys	s Ala	Phe	Thr	Phe	Ser	
				565					570)				575	ı		
	GAG	AAG	CCC	GTG	111	GTC	ТАТ	. GGG	СТС	сто	AC(GAA	CCC	GGC	GAC	AAG	1956
50	Glu	Lys	Pro	Val	Phe	. Val	Туг	Giy	Leu	Leu	Th:	Glu	Pro	Gly	Glu	Lys	
			580)				585	i				590)			

	AGC	ACG	GTC	TTC	ACC	ATC	TTT	GGC	TCC	AAC	AAG	GAG	AAC	GTG	CAC	AGC	2004
5	Ser	Thr	Val	Phe	Thr	He	Phe	Gly	Ser	Asn	Lys	Glu	Asn	Val	His	Ser	
		595					600					605					
	TGG	CTC	ATC	СТС	CAG	GTC	AAT	GCC	ACA	GAC	GCC	CTG	GGG	GTT	ССТ	TGC	2052
10	Trp	Leu	Ile	Leu	Gln	Val	Asn	Ala	Thr	Asp	Ala	Leu	Gly	Val	Pro	Cys	
	610					615					620					625	
	ACA	GAG	AAC	GAC	TAC	AAG	CTC	TGG	TCA	CCA	TCT	GAT	GAG	CGG	GGG	AAT	2100
15	Thr	Glu	Asn	Asp	Tyr	Lys	Leu	Trp	Ser	Pro	Ser	Asp	Glu	Arg	Gly	Asn	
•					630					635					640		
	GAG	TGT	TTG	CTT	GGA	CAC	AAG	ACT	GTT	TTC	AAA	CGG	AGG	ACC	CCG	CAC	2148
20	Glu	Cys	Leu	Leu	Gly	His	Lys	Thr	Val	Phe	Lys	Arg	Arg	Thr	Pro	His	
				645					650					655			
25	GCC	ACA	TGC	TTT	AAC	GGA	GAA	GAC	TTT	GAC	AGG	CCG	GTG	GTT	GTG	TCC	2196
	Ala	Thr	Cys	Phe	Asn	Gly	Glu	Asp	Phe	Asp	Arg	Pro	Val	Val	Val	Ser	
			660					665					670				
30	AAC	TGC	TCC	TGC	ACC	CGG	ĠAG	GAC	TAT	GAG	TGT	GAC	TTT	GGC	TTC	CGG	2244
	Asn	Cys	Ser	Cys	Thr	Arg	Glu	Asp	Туг	Glu	Cys	Asp	Phe	Gly	Phe	Arg	
		675					680					685					
35	ATG	AGT	GAA	GAC	TTG	GCA	TTA	GAG	GTG	TGT	GTT	CCA	GAT	CCA	GGA	TTT	2292
	Met	Ser	Glu	Asp	Leu	Ala	Leu	Glu	Val	Cys	Val	Pro	Asp	Pro	Gly	Phe	
	690					695					700					705	
	TCT	GGA	AAG	TCC	TCC	CCT	CCA	GTG	ССТ	TGT	CCC	GTG	GGC	тст	ACG	TAC	2340
	Ser	Gly	Lys	Ser	Ser	Pro	Pro	Val	Pro	Cys	Pro	Val	Gly	Ser	Thr	Tyr	
15					710					715					720		
	AGG	CGA	TCA	AGA	GGC	TAC	CGG	AAG	ATT	тст	GGG	GAC	ACC	TGT	AGT	GGA	2388
	Arg	Arg	Ser	Arg	Gly	Tyr	Arg	Lys	Ile	Ser	Gly	Asp	Thr	Cys	Ser	Gly	
50				725					730					735			
	GGA .	GAT	GTT	GAG	GCA	CGG	CTA	GAA	GGA	GAG	CTG	GTC	CCC	ፐርፕ	CCC	СТG	2436

	Gly	Asp	Val	Glu	Ala	Arg	Leu	Glu	Gly	Glu	Leu	Val	Pro	Cys	Pro	Leu	
_			740					745					750				
5	GCA	GAA	GAG	AAC	GAG	TTC	ATC	CTG	TAC	GCC	ACG	CGC	AAG	TCC	ATC	CAC	2484
	Ala	Glu	Glu	Asn	Glu	Phe	He	Leu	Туг	Ala	Thr	Arg	Lys	Ser	He	His	
10		755					760					765					
	CGC	TAT	GAC	CTG	GCT	TCC	GGA	ACC	ACG	GAG	CAG	TTG	CCC	СТС	ACT	GGG	2532
	Arg	Туг	Asp	Leu	Ala	Ser	Gly	Thr	Thr	Glu	Gln	Leu	Pro	Leu	Thr	Gly	
15	770					775					780					785	
	TTC	CGG	GCA	GCA	GTG	GCC	CTG	GAC	TTT	GAC	TAT	GAG	CAC	AAC	TGC	CTG	2580
	Leu	Arg	Ala	Ala	Val	Ala	Leu	Asp	Phe	Asp	Туг	Glu	His	Asn	Cys	Leu	
20					790					795					800		
	TAT	TGG	тст	GAC	CTG	GCC	TTG	GAC	GTC	ATC	CAG	CGC	СТС	TGT	TTG	AAC	2628
25	Tyr	Тгр	Ser	Asp	Leu	Ala	Leu	Asp	Val	lle	Gln	Arg	Leu	Cys	Leu	Asn	
				805					810					815			
	GGG	AGT	ACA	GGA	CAA	GAG	GTG	ATC	ATC	AAC	TCT	GAC	СТG	GAG	ACG	GTA	2676
30	Gly	Ser	Thr	Gly	Gln	Glu	Val	Ile	lle	Asn	Ser	Asp	Leu	Glu	Thr	Val	
			820					825					830				
	GAA	GCT	TTG	GCT	TTT	GAA	CCC	СТС	AGC	CAA	ATT	CTT	TAC	TGG	GTG	GAC	2724
35	Glu	۸la	Leu	Ala	Phe	Glu	Pro	Leu	Ser	Gln	Leu	Leu	Tyr	Trp	Val	Asp	
		835					840					845					
40	GCA	GGC	TTT	AAA	AAG	ATC	GAG	GTA	GCC	ААТ	CCA	GAT	GGT	GAC	TTC	CGA	2772
40	Ala	Gly	Phe	Lys	Lys	Ile	Glu	Val	Ala	Asn	Pro	Asp	Gly	Asp	Phe	Arg	
	850					855					860					865	
45	СТС	ACC	GTC	GTC	TAA	TCC	TCG	GTG	CTG	GAT	CGG	CCC	CGG	GCC	CTG	GTC	2820
	Leu	Thr	Val	Val	Asn	Ser	Ser	Val	Leu	Asp	Arg	Pro	Arg	Ala	Leu	Val	
					870					875					880		
50	CTT	GTG	CCC	CAA	GAA	GGG	ATC	ATG	TTC	TGG	ACC	GAC	TGG	GGA	GAC	CTG	2868
	Leu	Val	Pro	Gln	Glu	Gly	lle	Met	Phe	Trp	Thr	Asp	Тгр	Gly	Asp	Leu	

				885					890					895	•		
-	AAG	ССТ	GGG	ATT	ТАТ	CGG	AGC	AAC	ATG	GAC	GGA	тст	GCC	GCC	TAT	CGC	2916
5	Lys	Pro	Gly	lle	Туг	Arg	Ser	Asn	Met	Asp	Gly	Ser	Ala	Ala	Туг	Arg	
			900					905					910				
10	CTC	GTG	TCG	GAG	GAT	GTG	AAG	TGG	ccc	AAT	GGC	ATT	TCC	GTG	GAC	GAT	2964
	Leu	Val	Ser	Glu	Asp	Val	Lys	Trp	Pro	Asn	Gly	Пе	Ser	Val	Asp	Asp	
		915					920					925					
15	CAG	TGG	ATC	TAC	TGG	ACG	GAT	GCC	TAC	CTG	GÀC	TGC	АТТ	GAG	CGC	ATC	3012
	Gln	Trp	lle	Туг	Тгр	Thr	Asp	Ala	Tyr	Leu	Asp	Cys	Ile	Glu	Arg	Ile	
	930					935					940					945	
20	ACG	TTC	AGC	GGC	CAG	CAG	CGC	TCC	GTC	ATC	СТС	GAC	AGA	СТС	CCG	CAC	3060
	Thr	Phe	Ser	Gly	Gln	Gln	Arg	Ser	Val	He	Leu	Asp	Arg	Leu	Pro	His	
25					950					955					960		
,	CCC	TAT	GCC	ATT	GCT	GTC	TTT	AAG	ААТ	GAG	АТТ	TAC	TGG	GAT	GAC	TGG	3108
	Pro	Tyr	Ala	Ile	Ala	Val	Phe	Lys	Asn	Glu	He	Tyr	Trp	Asp	Asp	Trp	
30				965					970					975			
•	TCA	CAG	CTC	AGC	ATA	TTC	CGA	GCT	TCT	AAG	TAC	AGC	GGG	TCC	CAG	ATG	3156
	Ser	Gln	Leu	Ser	Ιlе	Phe	Arg	Ala	Ser	Lys	Туг	Ser	Gly	Ser	Gln	Met	
35			980					985					990				
	GAG	АТТ	СТС	GCC	AGC	CAG	CTC	ACG	GGG	CTG	ATG	GAC	ATG	AAG	ATC	TTC	3204
40	Glu	He	Leu	Ala	Ser	Gln	Leu	Thr	Gly	Leu	Met	Asp	Met	Lys	He	Phe	
••		995					100	0				100	5				
	TAC	AAG	GGG	AAG	AAC	ACA	GGA	AGC	AAT	GCG	TGT	GTA	CCC	AGG	CCG	TGC	3252
4 5	Tyr	Lys	Gly	Lys	Asn	Thr	Gly	Ser	Asn	Ala	Cys	Val	Pro	Arg	Pro	Cys	
	1016	0				101	5				102	0				1025	
	AGC	CTG	CTG	TGC	CTG	CCC	AGA	GCC	AAC	AAC	AGC	AAA	AGC	TGC	AGG	TGT	3300
50	Ser	Leu	Leu	Cys	Leu	Pro	Arg	Ala	Asn	Asn	Ser	Lys	Ser	Cys	Arg	Cys	
					103	0				103	5				104	0	

	CCA	GAT	GGC	GTG	GCC	AGC	AGT	GTC	CTC	CCT	TCC	GGG	GAC	CTG	ATG	TGT	3348
_	Pro	Asp	Gly	Val	Ala	Ser	Ser	Val	Leu	Pro	Ser	Gly	qaA	Leu	Met	Cys	
5				1045	5				1050)				1055	ō		
	GAC	TGC	CCT	AAG	GGC	TAC	GAG	CTG	AAG	AAC	AAC	ACG	TGT	GTC	AAA	GAA	3396
10	Asp	Cys	Рго	Lys	Gly	Туг	Glu	Leu	Lys	Asn	Asn	Thr	Cys	Val	Lys	Glu	
			1060)				1065	5				1070)			
	GAA	GAC	ACC	TGT	CTG	CGC	AAC	CAG	TAC	CGC	TGC	AGC	AAC	GGG	AAC	TGC	3444
15	Glu	Asp	Thr	Cys	Leu	Arg	Asn	Gln	Туг	Arg	Cys	Ser	Asn	Gly	Asn	Cys	
		1079	5				1080)			•	1089	5				
	ATC	AAC	AGC	ATC	TGG	TGG	TGC	GAT	TTC	GAC	AAC	GAC	TGC	GGA	GAC	ATG	3492
20	He	Asn	Ser	Ile	Trp	Trp	Cys	Asp	Phe	Asp	Asn	Asp	Cys	Gly	Asp	Met	
	1090)				1095	5				1100)				1105	
25	AGC	GAC	GAG	AAG	AAC	TGC	CCT	ACC	ACC	ATC	TGC	GAC	CTG	GAC	ACC	CAG	3540
	Ser	Asp	Glu	Lys	Asn	Cys	Pro	Thr	Thr	lle	Cys	Asp	Leu	Asp	Thr	Gln	
					1110)				1113	5				1120)	
30	TTC	CGT	тGC	CAG			GGG	ACG	TGC			стс	TCC	TAC		O TGT	3588
30					GAG	тст				ATC	CCG					TGT	3588
30					GAG Glu	тст				ATC Ile	CCG				AAA Lys	TGT	3588
30 35	Phe	Arg	Cys	Gln 112	GAG Glu 5	TCT Ser	Gly	Thr	Cys	ATC Ile	CCG Pro	Leu	Ser	Tyr 113	AAA Lys 5	TGT	3588 3636
	Phe	Ar g	Cys GAG	Gln 1125 GAT	GAG Glu 5 GAC	TCT Ser TGT	G1 y GGG	Thr	Cys 1130 AAC	ATC Ile O AGT	CCG Pro	Leu GAA	Ser AGG	Tyr 113 CAC	AAA Lys 5	TGT Cys	
35	Phe	Ar g	Cys GAG	Gln 1129 GAT Asp	GAG Glu 5 GAC	TCT Ser TGT	G1 y GGG	Thr	Cys 1130 AAC Asn	ATC Ile O AGT	CCG Pro	Leu GAA	Ser AGG	Tyr 113 CAC His	AAA Lys 5 TGT	TGT Cys	
	Phe GAC Asp	Arg CTC Leu	Cys GAG Glu 114	Gln 1129 GAT Asp	GAG Glu 5 GAC Asp	TCT Ser TGT Cys	Gly GGG Gly	Thr GAC Asp	Cys 1130 AAC Asn 5	ATC Ile O AGT Ser	CCG Pro GAC Asp	Leu GAA Glu	Ser AGG Arg 115	Tyr 113 CAC His	AAA Lys 5 TGT Cys	TGT Cys	
35	Phe GAC Asp	Arg CTC Leu CAC	Cys GAG Glu 114 CAG	Gln 112: GAT Asp 0 TGC	GAG Glu 5 GAC Asp	TCT Ser TGT Cys	Gly GGG Gly GAC	Thr GAC Asp 114 GAA	Cys 1130 AAC Asn 5	ATC Ile O AGT Ser	CCG Pro GAC Asp	GAA Glu AGC	Ser AGG Arg 115 TCG	Tyr 113 CAC His 0 GGC	AAA Lys 5 TGT Cys	TGT Cys GAA Glu	3636
35	Phe GAC Asp	Arg CTC Leu CAC	Cys GAG Glu 1144 CAG Gln	Gln 112: GAT Asp 0 TGC	GAG Glu 5 GAC Asp	TCT Ser TGT Cys	Gly GGG Gly GAC	GAC Asp 114 GAA Glu	Cys 1130 AAC Asn 5	ATC Ile O AGT Ser	CCG Pro GAC Asp	GAA Glu AGC	AGG Arg 115 TCG Ser	Tyr 113 CAC His 0 GGC	AAA Lys 5 TGT Cys	TGT Cys GAA Glu TGC	3636
35 40	Phe GAC Asp ATG Met	CTC Leu CAC His	GAG Glu 1144 CAG Gln 5	Gln 1129 GAT Asp O TGC Cys	GAG Glu 5 GAC Asp CGG Arg	TCT Ser TGT Cys AGC Ser	Gly GGG Gly GAC Asp 116	GAC Asp 114 GAA Glu	Cys 1130 AAC Asn 5 TAC	ATC Ile O AGT Ser AAC Asn	CCG Pro GAC Asp TGC Cys	GAA Glu AGC Ser 116	AGG Arg 115 TCG Ser	Tyr 113 CAC His O GGC	AAA Lys 5 TGT Cys ATG Met	TGT Cys GAA Glu TGC	3636 3684
35 40	Phe GAC Asp ATG Met	CTC Leu CAC His 115 CGC	GAG Glu 1144 CAG Gln 5	Gln 1125 GAT Asp 0 TGC Cys	GAG Glu GAC Asp CGG Arg	TCT Ser TGT Cys AGC Ser GTG	GGG Gly GAC Asp 116	GAC ASP 114 GAA Glu 0 GAC	Cys 1130 AAC Asn 5 TAC Tyr	ATC Ile O AGT Ser AAC ASn	CCG Pro GAC Asp TGC Cys	GAAA Glu AGC Ser 116 GAC	AGG Arg 115 TCG Ser 5	Tyr 113 CAC His 0 GGC Gly	AAA Lys 5 TGT Cys ATG Met	TGT Cys GAA Glu TGC Cys	3636
35 40	Phe GAC Asp ATG Met	CTC Leu CAC His 115 CGC	GAG Glu 1144 CAG Gln 5	Gln 1125 GAT Asp 0 TGC Cys	GAG Glu GAC Asp CGG Arg	TCT Ser TGT Cys AGC Ser GTG	GGG Gly GAC Asp 116 TGC Cys	GAC ASP 114 GAA Glu 0 GAC	Cys 1130 AAC Asn 5 TAC Tyr	ATC Ile O AGT Ser AAC ASn	CCG Pro GAC Asp TGC Cys	GAAA Glu AGC Ser 116 GAC Asp	AGG Arg 115 TCG Ser 5	Tyr 113 CAC His 0 GGC Gly	AAA Lys 5 TGT Cys ATG Met	TGT Cys GAA Glu TGC Cys	3636 3684

	Ser	Asp	Glu	Ala	Asn	Cys	Thr	Ala	Ile	Туг	His	Thr	Cys	Glu	Ala	Ser	
5					119	00				119	95				120	0	
	AAC	TTC	CAG	TGC	CGC	AAC	GGG	CAC	TGC	ATC	ccc	CAG	CGO	TGG	GCG	TGT	3828
	Asn	Phe	Gln	Cys	Arg	Asn	Gly	His	Cys	He	Pro	Gln	Arg	Trp	Ala	Cys	
10 -				120	5				121	0				121	5		
	GAC	GGC	GAC	GCC	GAC	TGC	CAG	GAT	GGC	TCT	GAT	GAG	GAT	CCA	GCC	AAC	3876
	Asp	Gly	Asp	Ala	Asp	Cys	Gln	Asp	Gly	Ser	Asp	Glu	Asp	Pro	Ala	Asn	
15			122	0				122	5				123	0			
	TGT	GAG	AAG	AAG	TGC	AAC	GGC	TTC	CGC	TGC	CCG	AAC	GGC	ACC	TGC	ATT	3924
20	Cys	Glu	Lys	Lys	Cys	Asn	Gly	Phe	Arg	Cys	Pro	Asn	Gly	Thr	Cys	Ile	
		123	5				124	0				124	5				
	CCC	TCC	ACC	AAG	CAC	TGT	GAC	GGC	CTG	CAC	GAT	TGC	TCG	GAC	GGC	TCC	3972
25	Pro	Ser	Thr	Lys	His	Cys	Asp	Gly	Leu	His	Asp	Cys	Ser	Asp	Gly	Ser	
	1250)				1255	õ				126	0				1265	
	GAC	GAG	CAG	CAC	TGC	GAG	ccc	CTG	TGT	ACA	CGG	TTC	ATG	GAC	TTC	GTG	4020
30	Asp	Glu	Gln	His	Cys	Glu	Pro	Leu	Cys	Thr	Arg	Phe	Met	Asp	Phe	Val	
					1270)		,		1275	5				1280)	
35	TGT	AAG	AAC	CGC	CAG	CAG	TGC	СТС	TTC	CAC	TCC	ATG	GTG	TGC	GAT	GGG	4068
33	Cys	Lys	Asn	Arg	Gln	Gln	Cys	Leu	Phe	His	Ser	Met	Val	Cys	Asp	Gly	
				1285	j				1290)				1295	5		
40	ATC	ATC	CAG	TGC	CGT	GAC	GGC	тсс	GAC	GAG	GAC	CCA	GCC	ттт	GCA	GGA	4116
	Πe	lle	Gln	Cys	Arg	Asp	Gly	Ser	Asp	Glu	Asp	Pro	Ala	Phe	Ala	Gly	
			1300)				1305	i				1310)			
45	TGC	TCC	CGA	GAC	CCC	GAG	TTC	CAC	AAG	GTG	TGC	GAT	GAG	TTC	GGC	TTC	4164
	Cys	Ser	Arg	Asp	Pro	Glu	Phe	His	Lys	Val	Cys	Asp	Glu	Phe	Gly	Phe	
		1315	j				1320)				1325	5				
50	CAG	TGT	CAG	AAC	GGC	GTG	TGC	ATC	AGC	TTG	ATC	TGG	AAG	TGC	GAC	GGG	4212
	Gln	Cys	Gln	Asn	Gly	Val	Cys	Ile	Ser	Leu	Ile	Trp	Lys	Cys	Asp	Gly	

	1330					1335					1340)			1345		
5	ATG	GAT	GAC	TGC	GGG	GAC	TAC	TCC	GAC	GAG	GCC	AAC	TGT	GAA	AAC	CCC	4260
5	Met	Asp	Asp	Cys	Gly	Asp	Tyr	Ser	Asp	Glu	Ala	Asn	Cys	Glu	Asn	Pro	
		1350								1355	5				1360)	
10	ACA	GAA	GCC	CCC	AAC	TGC	TCC	CGC	TAC	TTC	CAG	TTC	CGG	TGT	GAC	ΑΑΤ	4308
	Thr	Glu	Ala	Pro	Asn	Cys	Ser	Arg	Туг	Phe	Gln	Phe	Arg	Cys	Asp	Asn	
				1365	5				1370)				1375	5		
15	GGC	CAC	TGC	ATC	CCC	AAC	AGG	TGG	AAG	TGT	GAC	AGG	GAG	AAT	GAC	TGT	4356
	Gly	His	Cys	Ile	Pro	Asn	Arg	Trp	Lys	Cys	Asp	Arg	Glu	Asn	Asp	Cys	
	1380							1385	5				1390)			
20	GGG	GAC	TGG	TCC	GAC	GAG	AAG	GAC	TGT	GGA	GÁT	TCA	CAT	GTA	CTT	CCG	4404
	Gly	Asp	Trp	Ser	Asp	Glu	Lys	Asp	Cys	Gly	Asp	Ser	His	Val	Leu	Pro	
25	1395						1400)		1405							
	TCT	ACG	ACT	CCT	GCA	CCC	TCC	ACG	TCT	CTG	CCC	AAT	TAC	TAC	CGC	TGC	4452
	Ser	Thr	Thr	Pro	Ala	Pro	Ser	Thr	Cys	Leu	Pro	Asn	Tyr	Туг	Arg	Cys	
30	1410)				1415	5				1420)				1425	
	GGC	GGG	GGG	GCC	TGC	GTG	ATA	GAC	ACG	TGG	GTT	TGT	GAC	GGG	TAC	CGA	4500
	Gly	Gly	Gly	Ala	Cys	Val	He	Asp	Thr	Trp	Val	Cys	Asp	Gly	Tyr	Arg	
35					1430)				1435	5 .				1440)	
	GAT	TGC	GCA	GAT	GGA	TCC	GAC	GAG	GAA	GCC	TGC	CCC	TCG	CTC	CCC	AAT	4548
40	Asp	Cys	Ala	Asp	Gly	Ser	Asp	Glu	Glu	Ala	Cys	Pro	Ser	Leu	Pro	Asn	
				1445	5				1450)				145	5		
	GTC	ACT	GCC	ACC	TCC	TCC	CCC	TCC	CAG	CCT	GGA	CGA	TGC	GAC	CGA	TTT	4 596
45	Val	Thr	Ala	Thr	Ser	Ser	Pro	Ser	Gln	Pro	Gly	Arg	Cys	Asp	Arg	Phe	
			146	0				146	5				1470	0			
	GAG	TTT	GAG	TGC	CAC	CAG	CCA	AAG	AAG	TGC	ATC	CCT	AAC	TGG	AGA	CGC	4644
50	Glu	Phe	Glu	Cys	His	Gln	Pro	Lys	Lys	Cys	lle	Pro	Asn	Trp	Arg	Arg	
	1775						148	0				148	5				

	TGT	GAC	GGC	CAT	CAG	GAT	TGC	CAG	GAT	GGC	CAG	GAC	GAG	GCC	AAC	TGC	4692
5	Cys	Asp	Gly	His	Gln	Asp	Cys	Gln	Asp	Gly	Gln	Asp	Glu	Ala	Asn	Cys	
	1490)				1499	5				1500)				1505	
	CCC	ACT	CAC	AGC	ACC	TTG	ACC	TGC	ATG	AGC	TGG	GAG	TTC	AAG	TGT	GAG	4740
10	Pro	Thr	His	Ser	Thr	Leu	Thr	Cys	Met	Ser	Trp	Glu	Phe	Lys	Cys	Glu	
					1510)				1515	5				1520)	
	GAT	GGC	GAG	GCC	TGC	ATC	GTG	CTG	TCA	GAA	CGC	TGC	GAC	GGC	TTC	CTG	4788
15	Asp	Gly	Glu	Ala	Cys	He	Val	Leu	Ser	Glu	Arg	Cys	Asp	Gly	Phe	Leu	
				1525					1530)				1535	5		
	GAC	TGC	TCA	GAT	GAG	AGC	GAC	GAG	AAG	GCC	TGC	AGT	GAT	GAG	TTA	ACT	4836
20	Asp	Cys	Ser	Asp	Glu	Ser	Asp	Glu	Lys	Ala	Cys	Ser	Asp	Glu	Leu	Thr	
			1540)				1549	5				1550)			
25	GTA	TAC	AAA	GTA	CAG	AAT	CTT	CAG	TGG	ACA	GCT	GAC	TTC	тст	GGG	ΑΑΤ	4884
	Val	Туг	Lys	Val	Gln	Asn	Leu	Gln	Trp	Thr	Ala	Asp	Phe	Ser	Gly	Asn	
		1555	5				1560)				1565	5				
30	GTC	ACT	TTG	ACC	TGG	ATG	CGG	CCC	AAA	AAA	ATG	CCC	TCT	GCT	GCT	TGT	4932
30				ACC Thr													4932
30		Thr					Arg					Pro					4932
30 35	Val 1570	Thr	Leu		Trp	Met 1575	Arg	Pro	Lys	Lys	Met 1580	Pro)	Ser	Ala	Ala	Cys 1585	4932 4980
	Val 1570 GTA	Thr) TAC	Leu AAC	Thr	Trp TAC	Met 1575 TAT	Arg S AGA	Pro GTT	Lys GTT	Lys GGA	Met 1580 GAG	Pro) AGC	Ser ATA	Ala TGG	Ala AAG	Cys 1585 ACT	
35	Val 1570 GTA	Thr) TAC	Leu AAC	Thr GTG	Trp TAC	Met 1575 TAT Tyr	Arg S AGA	Pro GTT	Lys GTT	Lys GGA	Met 1580 GAG Glu	Pro) AGC	Ser ATA	Ala TGG	Ala AAG	Cys 1585 ACT Thr	
	Val 1570 GTA Val	Thr) TAC Tyr	Leu AAC Asn	Thr GTG	TAC Tyr 1590	Met 1575 TAT Tyr	Arg AGA Arg	Pro GTT Val	Lys GTT Val	Lys GGA Gly 1595	Met 1580 GAG Glu	Pro) AGC Ser	Ser ATA Ile	Ala TGG Trp	Ala AAG Lys 1600	Cys 1585 ACT Thr	
35	Val 1570 GTA Val	Thr TAC Tyr GAG	Leu AAC Asn ACT	Thr GTG Val	TAC Tyr 1590 AGC	Met 1575 TAT Tyr	Arg AGA Arg	Pro GTT Val	Lys GTT Val	Lys GGA Gly 1595 ACT	Met 1580 GAG Glu OTA	Pro AGC Ser	Ser ATA Ile	Ala TGG Trp GTG	Ala AAG Lys 1600 TTG	Cys 1585 ACT Thr	4980
35	Val 1570 GTA Val	Thr TAC Tyr GAG	Leu AAC Asn ACT	Thr GTG Val	TAC Tyr 1590 AGC Ser	Met 1575 TAT Tyr	Arg AGA Arg	Pro GTT Val	Lys GTT Val	Lys GGA Gly 1595 ACT Thr	Met 1580 GAG Glu OTA	Pro AGC Ser	Ser ATA Ile	Ala TGG Trp GTG	Ala AAG Lys 1600 TTG Leu	Cys 1585 ACT Thr	4980
35 10	Val 1570 GTA Val CTG Leu	Thr TAC Tyr GAG Glu	Leu AAC Asn ACT Thr	Thr GTG Val CAC	TAC Tyr 1590 AGC Ser	Met 1575 TAT Tyr) AAT	Arg AGA Arg AAG Lys	Pro GTT Val ACA Thr	Lys GTT Val AAC Asn 1610	GGA Gly 1595 ACT Thr	Met 1580 GAG Glu 5 GTA Val	Pro AGC Ser TTA Leu	Ser ATA Ile AAA Lys	TGG Trp GTG Val	Ala AAG Lys 1600 TTG Leu	Cys 1585 ACT Thr AAA Lys	4980
35 10	Val 1570 GTA Val CTG Leu	Thr TAC Tyr GAG Glu GAT	Leu AAC Asn ACT Thr	Thr GTG Val CAC His	TAC Tyr 1590 AGC Ser 5	Met 1575 TAT Tyr AAT Asn CAG	Arg AGA Arg AAG Lys	Pro GTT Val ACA Thr	CTT Val AAC Asn 1610 GTG	GGA Gly 1598 ACT Thr CAG	Met 1580 GAG Glu 5 GTA Val	Pro) AGC Ser TTA Leu CAG	Ser ATA Ile AAA Lys	TGG Trp GTG Val 1615	AAAG Lys 16000 TTG Leu AGC	Cys 1585 ACT Thr AAA Lys	4980 5028
35 10	Val 1570 GTA Val CTG Leu	Thr TAC Tyr GAG Glu GAT	Leu AAC Asn ACT Thr	Thr GTG Val CAC His 1605 ACC Thr	TAC Tyr 1590 AGC Ser 5	Met 1575 TAT Tyr AAT Asn CAG	Arg AGA Arg AAG Lys	Pro GTT Val ACA Thr	CTT Val AAC Asn 1610 GTG Val	GGA Gly 1598 ACT Thr CAG	Met 1580 GAG Glu 5 GTA Val	Pro) AGC Ser TTA Leu CAG	Ser ATA Ile AAA Lys	TGG Trp GTG Val 1615 CTG Leu	AAAG Lys 16000 TTG Leu AGC	Cys 1585 ACT Thr AAA Lys	4980 5028
55 50	Val 1570 GTA Val CTG Leu CCA	Thr TAC Tyr GAG Glu GAT Asp	AAC Asn ACT Thr ACC Thr 1620	Thr GTG Val CAC His 1605 ACC Thr	Trp TAC Tyr 1590 AGC Ser TAC Tyr	Met 1575 TAT Tyr AAT Asn CAG	Arg AGA Arg AAG Lys GTT Val	Pro GTT Val ACA Thr AAA Lys 1623	CTT Val AAC Asn 1610 GTG Val	GGA Gly 1598 ACT Thr CAG	Met 1580 GAG Glu 5 GTA Val GTT Val	Pro CAG CAG GIn	Ser ATA Ile AAA Lys TGC Cys 1630	TGG Trp GTG Val 1615 CTG Leu	AAAG Lys 16000 TTG Leu AGC Ser	Cys 1585 ACT Thr AAA Lys AAG Lys	4980 5028

	Val	His	Asn	Thr	Asn	Asp	Phe	Val	Thr	Leu	Arg	Thr	Pro	Glu	Gly	Leu	
_		1635	ō				1640)				1645	;				
5	CCA	GAC	GCC	CCT	CAG	AAC	CTC	CAG	CTG	TCG	CTC	CAC	GGG	GAA	GAG	GAA	5172
	Pro	Asp	Ala	Pro	Gln	Asn	Leu	Gln	Leu	Ser	Leu	His	Gly	Glu	Glu	Glu	
10	1656)				1655	5				1660)				1665	
	GGT	GTG	ATT	GTG	GGC	CAC	TGG	AGC	CCT	CCC	ACC	CAC	ACC	CAC	GGC	CTC	5220
	Gly	Val	Ile	Val	Gly	His	Trp	Ser	Pro	Pro	Thr	His	Thr	His	Gly	Leu	
15					1670)				1675	5				1680)	
	ATT	CGC	GAA	TAC	ATT	GTA	GAG	TAT	AGC	AGG	AGT	GGT	TCC	AAG	GTG	TGG	5268
	Ile	Arg	Glu	Tyr	Ile	Val	Glu	Tyr	Ser	Arg	Ser	Gly	Ser	Lys	Val	Trp	
20				168	5				1696)				169	5		
	ACT	TCA	GAA	AGG	GCT	GCT	AGT	AAC	TTT	ACA	GAA	ATA	AAG	AAC	TTG	TTG	5316
25	Thr	Ser	Glu	Arg	Ala	Ala	Ser	Asn	Phe	Thr	Glu	He	Lys	Asn	Leu	Leu	
			170	0				170	5				171	0			
	GTC	AAC	ACC	СТG	TAC	ACC	GTC	AGA	GTG	GCT	GCG	GTG	ACG	AGT	, CGT	GGG	5364
30	Val	Asn	Thr	Leu	Туг	Thr	Val	Arg	Val	Ala	Ala	Val	Thr	Ser	Arg	Gly	
		171	5				172	:0				172	5				
	ATA	A GGA	AAC	TGC	G AGC	GA1	г тсс	CAAA	TC0	AT1	r ac	CACC	C GTO	G AA	A GG	A AAA	5412
35	П	e Gly	/ Asn	Trp	Ser	Asp	Sei	Lys	s Sei	- 116	e Th	r Thi	· Va	l Ly	s Gl	y Lys	
	173	30				173	35				17	40				1745	j.
10	GC	G ATO	ccc	CC/	CCA	AA'	TA 1	C CAC	C AT	r GA	CAA	C TAC	C GA	T GA	A AA'	T TCC	5460
40	Al	a Ile	e Pro	Pro) Pro	Ası	n Il	e His	s II	e Ası	p As	n Ty	r As	p G1	u As	n Ser	
					175	50				17	55				17	60	
45	СТ	G AG	r TT	r ac	CTO	G AC	C GT	G GA	T GG	G AA	C AT	C AA	G GT	G AA	T GG	С ТАТ	5508
	Le	u Se	r Phe	e Th	r Lei	ı Th	r Va	l As	p Gl	y As	n II	e Ly	s Va	l As	n Gl	y Tyr	
				17	65				17	70				17	75		
50	GT	G GT	G AAG	с ст	т тт	C TG	G GC	а тт	T GA	C AC	C CA	C AA	A CA	A GA	G AA	G AAA	5556
	۷a	l Va	l As	n Le	u Ph	e Tr	p Al	a Ph	e As	p Th	r Hi	s Ly	s Gl	n Gl	u Ly	s Lys	

	1780					1785						1790					
5	ACC	ATG	AAC	TTC	CAA	GGG	AGC	TCA	GTG	TCC	CAC	AAA	GTT	GGC	AAT	CTG	5604
	Thr	Met	Asn	Phe	GIn	Gly	Ser	Ser	Val	Ser	His	Lys	Val	Gly	Asn	Leu	
		179	5				180	0				1805	5				
10	ACA	GCA	CAG	ACG	GCC	TAT	GAG	ATT	TCC	GCC	TGG	GCC	AAG	ACT	GAC	TTG	5652
	Thr	Ala	Gln	Thr	Ala	Туг	Glu	Ile	Ser	Ala	Trp	Ala	Lys	Thr	Asp	Leu	
	1810)				1819	5				1820)		-		1825	
15	GGC	GAT	AGT	ССТ	CTG	TCA	TTT	GAG	CAT	GTC	ACG	ACC	AGA	GGG	GTT	CGC	5700
	Gly	Asp	Ser	Pro	Leu	Ser	Phe	Glu	His	Val	Thr	Thr	Arg	Gly	Val	Arg	
					1830)				1835	5			•	1840)	
20	CCA	CCT	GCT	ССТ	AGC	СТС	AAG	GCC	AGG	GCT	ATC	AAT	CAG	АСТ	GCA	GTG	5748
	Pro	Pro	Ala	Pro	Ser	Leu	Lys	Ala	Arg	Ala	Πle	Asn	Gln	Thr	Ala	Val	
25				1845	5				1850)				1859	5		
	GAA	TGC	ACC	TGG	ACA	GGC	CCC	AGG	AAT	GTG	GŤG	TAT	GGC	ATT	TTC	ТАТ	5796
	Glu	Cys	Thr	Trp	Thr	Gly	Pro	Arg	Asn	Val	Val	Tyr	Gly	Пе	Phe	Tyr	
30			1860)				1865	5				1870)			
	GCC	ACA	TCC	TTC	CTG	GAC	СТС	TAC	CGC	AAC	CCA	AGC	AGC	CTG	ACC	ACG	5844
	Ala	Thr	Ser	Phe	Leu	Asp	Leu	Туг	Arg	Asn	Pro	Ser	Ser	Leu	Thr	Thr	
35		1875	5				1880)				1885	5				
	CCG	CTG	CAC	AAC	GCA	ACC	GTG	CTC	GTC	GGT	AAG	GAT	GAG	CAG	TAT	CTG	5892
40	Pro	Leu	His	Asn	Ala	Thr	Val	Leu	Val	Gly	Lys	Asp	Glu	Gln	Tyr	Leu	
10	1890)				1895	5				1900)				1905	
	TTT	CTG	GTC	CGG	GTG	GTG	ATG	CCC	TAC	CAA	GGG	CCG	TCC	TCG	GAC	TAC	5940
15	Phe	Leu	Val	Arg	Val	Val	Met	Pro	Туг	Gln	Gly	Pro	Ser	Ser	Asp	Tyr	
					1910)				1915	5				1920		
	GTG	GTC	GTG	AAG	ATG	ATC	CCG	GAC	AGC	AGG	CTT	ССТ	ccc	CGG	CAC	CTG	5988
ю	Val	Val	Val	Lys	Met	Ιlе	Pro	Asp	Ser	Arg	Leu	Pro	Pro	Arg	His	Leu	
				1925	j				1930)				1935	5		

	CAT	GCC	GTT	CAC	ACC	GGC	AAG	ACC	TCG	GCC	GTC	ATC	AAG	TGG	GAG	TCG	6036
5	His	Ala	Val	His	Thr	Gly	Lys	Thr	Ser	Ala	Val	Пe	Lys	Trp	Glu	Ser	
•			194	0				194	5				195	0			
	CCC	TAC	GAC	TCT	ССТ	GAC	CAG	GAC	CTG	TTC	TAT	GCG	ATC	GCA	GTT	AAA	6084
10	Pro	Туг	Asp	Ser	Pro	Asp	G1 n	Asp	Leu	Phe	Туг	Ala	Ile	Ala	Val	Lys	
		195	5				196	0				196	วี				
	GAT	CTG	ATA	CGA	AAG	ACG	GAC	CGG	AGC	TAC	AAA	GTC	AAG	TCC	CGC	AAC	6132
15	Asp	Leu	Ile	Arg	Lys	Thr	Asp	Arg	Ser	Tyr	Lys	Val	Lys	Ser	Arg	Asn	
	1970					1975					1980)				1985	
	AGC	ACC	GTG	GAG	TAC	ACC	СТG	AGC	AAG	CTG	GAG	CCC	GGA	GGG	AAA	TAC	6180
20	Ser	Thr	Val	Glu	Туг	Thr	Leu	Ser	Lys	Leu	Glu	Pro	Gly	Gly	Lys	Tyr	
					1990)				1995	5				2000)	
25	CAC	GTC	ATT	GTG	CAG	CTG	GGG	AAC	ATG	AGC	AAA	GAT	GCC	AGT	GTG	AAG	6228
	His	Val	Ile	Val	Gln	Leu	Gly	Asn	Met	Ser	Lys	Asp	Ala	Ser	Val	Lys	
				2005	5				2010)				2015	5		
30	ATC	ACC	ACC			TTA	тсс	GCA			GCC	TTA	AAA			ACA	6276
30				GTT	TCG			GCA Ala	CCC	GAT				ATC	АТА		6276
30				GTT Val	TCG				CCC Pro	GAT				ATC Ile	АТА		6276
35	Ile	Thr	Thr 2020	GTT Val	TCG Ser	Leu	Ser	Ala	CCC Pro	GAT Asp	Ala	Leu	Lys 2030	ATC Ile	ATA Ile	Thr	6276 6324
35	Ile	Thr AAT	Thr 2020 GAC	GTT Val) CAC	TCG Ser GTC	Leu CTT	Ser CTC	Ala 2025 TTC	CCC Pro 5	GAT Asp AAA	Ala AGT	Leu CTA	Lys 2030 GCT	ATC Ile) CTA	ATA lle AAG	Thr	
35	Ile GAA	Thr AAT	Thr 2020 GAC Asp	GTT Val) CAC	TCG Ser GTC	Leu CTT	Ser CTC	Ala 2025 TTC Phe	CCC Pro 5	GAT Asp AAA	Ala AGT	Leu CTA	Lys 2030 GCT Ala	ATC Ile) CTA	ATA lle AAG	Thr	
35	Ile GAA Glu	Thr AAT Asn 2035	Thr 2020 GAC Asp	GTT Val) CAC His	TCG Ser GTC Val	Leu CTT Leu	Ser CTC Leu 2040	Ala 2025 TTC Phe	CCC Pro TGG Trp	GAT Asp AAA Lys	Ala AGT Ser	CTA Leu 2045	Lys 2030 GCT Ala	ATC Ile) CTA Leu	ATA lle AAG Lys	Thr GAA Glu	
35	Ile GAA Glu AAG	Thr AAT Asn 2035 TAT	Thr 2020 GAC Asp	GTT Val CAC His	TCG Ser GTC Val	CTT Leu	Ser CTC Leu 2040 AGG	Ala 2025 TTC Phe	CCC Pro TGG Trp	GAT ASP AAA Lys GAG	Ala AGT Ser	CTA Leu 2045 CAC	Lys 2030 GCT Ala ATG	ATC Ile) CTA Leu	ATA lle AAG Lys	Thr GAA Glu AGC	6324
35	Ile GAA Glu AAG	Thr AAT Asn 2035 TAT Tyr	Thr 2020 GAC Asp	GTT Val CAC His	TCG Ser GTC Val	CTT Leu	Ser CTC Leu 2040 AGG Arg	Ala 2025 TTC Phe) GGC	CCC Pro TGG Trp	GAT ASP AAA Lys GAG	Ala AGT Ser	CTA Leu 2045 CAC His	Lys 2030 GCT Ala ATG	ATC Ile) CTA Leu	ATA lle AAG Lys	Thr GAA Glu AGC	6324
35 40	GAA Glu AAG Lys 2050	Thr AAT Asn 2035 TAT Tyr	Thr 2020 GAC Asp TTT Phe	GTT Val CAC His AAC	TCG Ser GTC Val GAA Glu	CTT Leu AGC Ser 2055	CTC Leu 2040 AGG Arg	Ala 2025 TTC Phe) GGC	CCC Pro 5 TGG Trp TAC	GAT Asp AAA Lys GAG Glu	Ala AGT Ser ATA Ile 2060	CTA Leu 2045 CAC His	Lys 2030 GCT Ala ATG Met	ATC Ile) CTA Leu TTT Phe	ATA Ile AAG Lys GAT Asp	Thr GAA Glu AGC Ser 2065	6324
35 40	GAA Glu AAG Lys 2050 GCC	AAT ASN 2035 TAT Tyr ATG	Thr 2020 GAC Asp TTT Phe	GTT Val CAC His AAC Asn	TCG Ser GTC Val GAA Glu	CTT Leu AGC Ser 2055 GCA	CTC Leu 2040 AGG Arg	Ala 2025 TTC Phe) GGC Gly	CCC Pro TGG Trp TAC Tyr	GAT ASP AAA Lys GAG Glu AAT	Ala AGT Ser ATA Ile 2060 ACT	CTA Leu 2045 CAC His	Lys 2030 GCT Ala ATG Met	ATC Ile) CTA Leu TTT Phe	ATA Ile AAG Lys GAT Asp	Thr GAA Glu AGC Ser 2065 TTT	632 4 6372
35 40	GAA Glu AAG Lys 2050 GCC	AAT ASN 2035 TAT Tyr ATG	Thr 2020 GAC Asp TTT Phe	GTT Val CAC His AAC Asn	TCG Ser GTC Val GAA Glu	CTT Leu AGC Ser 2055 GCA Ala	CTC Leu 2040 AGG Arg	Ala 2025 TTC Phe) GGC Gly	CCC Pro TGG Trp TAC Tyr	GAT ASP AAA Lys GAG Glu AAT	Ala AGT Ser ATA Ile 2060 ACT Thr	CTA Leu 2045 CAC His	Lys 2030 GCT Ala ATG Met	ATC Ile) CTA Leu TTT Phe	ATA Ile AAG Lys GAT Asp	Thr GAA Glu AGC Ser 2065 TTT Phe	632 4 6372

	Lys	lle	Ser	Asn	Leu	Lys	Met	Gly	His	Asn	Туг	Thr	Phe	Thr	Val	Gln	
5				2085	5				2090)				2095	5		
	GCA	CGA	TGC	CTT	TTG	GGC	AGC	CAG	ATC	TGC	GGG	GAG	CCT	GCC	GTG	CTA	6516
	Ala	Arg	Cys	Leu	Leu	Gly	Ser	Gln	Ile	Cys	Gly	Glu	Pro	Ala	Val	Leu	
10			2100)				2105	5				2110)			
	CTG	TAT	GAT	GAG	CTG	GGG	TCT	GGT	GGC	GAT	GCG	TCG	GCG	ATG	CAG	GCT	6564
15	Leu	Tyr	Asp	Glu	Leu	Gly	Ser	Gly	Gly	Asp	Ala	Ser	Ala	Met	Gln	Ala	
73		2115	5				2120)				2125	5				
	GCC	AGG	TCT	ACT	GAT	GTC	GCC	GCC	GTG	GTG	GTG	CCC	ATC	СТG	TTT	CTG	6612
20	Ala	Arg	Ser	Thr	Asp	Val	Ala	Ala	Val	Val	Val	Pro	He	Leu	Phe	Leu	
	2130)				2135	5				2140)				2145	
	ATA	CTG	CTG	AGC	CTG	GGG	GTC	GGG	TTT	GCC	ATC	CTG	TAC	ACG	AAG	CAT	6660
25	Пе	Leu	Leu	Ser	Leu	Gly	Val	Gly	Phe	Ala	Ile	Leu	Tyr	Thr	Lys	His	
					2150)				215	5				216	0	
	CGG	AGG	CTG	CAG	AGC	AGC	TTC	ACC	GCC	TTC	GCC	AAC	AGC	CAC	TAC	AGC	6708
30	Arg	Arg	Leu	Gln	Ser	Ser	Phe	Thr	Ala	Phe	Ala	Asn	Ser	His	Tyr	Ser	
				216	5				217	0				217	5		
	TCC	AGA	CTC	GGC	TCC	GCC	ATC	TTC	TCC	тст	GGG	GAT	GAC	TTG	GGG	GAG	6756
35	Ser	Arg	Leu	Gly	Ser	Ala	lle	Phe	Ser	Ser	Gly	Asp	Asp	Leu	Gly	Glu	
			218	0				218	5				219	0			
40	GAT	GAT	GAA	GAT	GCT	ССТ	ATG	ATC	ACT	GGA	TTT	TCG	GAC	GAC	GTC	ccc	6804
	Asp	Asp	Glu	Asp	Ala	Pro	Met	Ile	Thr	Gly	Phe	Ser	Asp	Asp	Val	Pro	
		219	5				220	0				220	5				
45	ATG	GTG	АТА	GCC	TGA	AAGA	GCT	TTCC	TCAC	TA G	AAAC	CAAA	τ GG	TGTA	ΑΑΤΑ		6856
	Met	Val	Ile	Ala													
	221	0			•												
50	TTT	татт	тgа	TAAA	.GATA	GT T	GATO	GTTT	'A TT	ኅ ፐልል	AAGA	TGC	ACTI	ƳGA	GTTC	СААТАТ	6916
	GTT	'ATTT	ATT	тато	GGCC	AA A	AACA	AAAG	C AA	AAAA	AAAA	AAA	.AA				6961

	Sequence ID No. 4	
5	Length of the Sequence: 300	
	Type: nucleic acid	
	Strandedness: double	
10	Topology: linear	
	Molecular type: cDNA to mRNA	
	Sequence:	
15	ATATCCACAT TGACAGCTAT GGTGAAAATT ATCTAAGCTT CACCCTGACC ATGGAGAGTG	60
	ATATCAAGGT GAATGGCTAT GTGGTGAACC TTTTCTGGGC ATTTGACACC CACAAGCAAG	120
20	AGAGGAGAAC TTTGAACTTC CGAGGAAGCA TATTGTCACA CAAAGTTGGC AATCTGACAG	180
	CTCATACATC CTATGAGATT TCTGCCTGGG CCAAGACTGA CTTGGGGGAT AGCCCTCTGG	240
	CATTTGAGCA TGTTATGACC AGAGGGGTTC GCCCACCTGC ACCTAGCCTC AAGGCCAAAG	300
25	Sequence ID No. 5	
	Length of the Sequence: 6642	
	Type: nucleic acid	
30	Strandedness: double	
•	Topology: linear	
35	Molecular type: cDNA to mRNA	
	Sequence:	
	ATGGCGACAC GGAGCAGCAG GAGGGAGTCG CGACTCCCGT TCCTATTCAC CCTGGTCGCA	60
40	CTGCTGCCGC CCGGAGCTCT CTGCGAAGTC TGGACGCAGA GGCTGCACGG CGGCAGCGCG	120
	CCCTTGCCCC AGGACCGGGG CTTCCTCGTG GTGCAGGGCG ACCCGCGCGA GCTGCGGCTG	180
45	TGGGCGCGC GGGATGCCAG GGGGGCGAGC CGCGCGGACG AGAAGCCGCT CCGGAGGAAA	240
43	CGGAGCGCTG CCCTGCAGCC CGAGCCCATC AAGGTGTACG GACAGGTTAG TCTGAATGAT	3 00
	TCCCACAATC AGATGGTGGT GCACTGGGCT GGAGAGAAAA GCAACGTGAT CGTGGCCTTG	360
50	GCCCGAGATA GCCTGGCATT GGCGAGGCCC AAGAGCAGTG ATGTGTACGT GTCTTACGAC	420
	TATGGAAAAT CATTCAAGAA AATTTCAGAC AAGTTAAACT TTGGCTTGGG AAATAGGAGT	480

	GAAGCTGTTA	TCGCCCAGTT	CTACCACAGC	CCTGCGGACA	ACAAGCGGTA	CATCTTTGCA	540
5	GACGCTTATG	CCCAGTACCT	CTGGATCACG	TTTGACTTCT	GCAACACTCT	TCAAGGCTTT	600
	TCCATCCCAT	TTCGGGCAGC	TGATCTCCTC	CTACACAGTA	AGGCCTCCAA	CCTTCTCTTG	660
	GGCTTTGACA	GGTCCCACCC	CAACAAGCAG	CTGTGGAAGT	CAGATGACTT	TGGCCAGACC	720
10	TGGATCATGA	TTCAGGAACA	TGTCAAGTCC	TTTTCTTGGG	GAATTGATCC	CTATGACAAA	780
	CCAAATACCA	TCTACATTGA	ACGACACGAA	CCCTCTGGCT	ACTCCACTGT	CTTCCGAAGT	840
	ACAGATTTCT	TCCAGTCCCG	GGAAAACCAG	GAAGTGATCC	TTGAGGAAGT	GAGAGATTTT	900
15	CAGCTTCGGG	ACAAGTACAT	GTTTGCTACA	AAGGTGGTGC	ATCTCTTGGG	CAGTGAACAG	960
	CAGTCTTCTG	TCCAGCTCTG	GGTCTCCTTT	GGCCGGAAGC	CCATGAGAGC	AGCCCAGTTT	1020
20	GTCACAAGAC	ATCCTATTAA	TGAATATTAC	ATCGCAGATG	CCTCCGAGGA	CCAGGTGTTT	1080
20	GTGTGTGTCA	GCCACAGTAA	CAACCGCACC	AATTTATACA	TCTCAGAGGC	AGAGGGGCTG	1140
	AAGTTCTCCC	TGTCCTTGGA	GAACGTGCTC	TATTACAGCC	CAGGAGGGC	CGGCAGTGAC	1200
25	ACCTTGGTGA	GGTATTTTGC	AAATGAACCA	TTTGCTGACT	TCCACCGAGT	GGAAGGATTG	1260
	CAAGGAGTCT	ACATTGCTAC	TCTGATTAAT	GGTTCTATGA	ATGAGGAGAA	CATGAGATCG	1320
	GTCATCACCT	TTGACAAAGG	GGGAACCTGG	GAGTTTCTTC	AGGCTCCAGC	CTTCACGGGA	1380
30	TATGGAGAGA	AAATCAATTG	TGAGCTTTCC	CAGGGCTGTT	CCCTTCATCT	GGCTCAGCGC	1440
	CTCAGTCAGC	TCCTCAACCT	CCAGCTCCGG	AGAATGCCCA	TCCTGTCCAA	GGAGTCGGCT	1500
25	CCAGGCCTCA	TCATCGCCAC	TGGCTCAGTG	GGAAAGAACT	TGGCTAGCAA	GACAAACGTG	1560
35	TACATCTCTA	GCAGTGCTGG	AGCCAGGTGG	CGAGAGGCAC	TTCCTGGACC	ТСАСТАСТАС	1620
	ACATGGGGAG	ACCACGGCGG	AATCATCACG	GCCATTGCCC	AGGGCATGGA	AACCAACGAG	1680
40	CTAAAATACA	GTACCAATGA	AGGGGAGACC	TGGAAAACAT	TCATCTTCTC	TGAGAAGCCA	1740
	GTGTTTGTGT	ATGGCCTCCT	CACAGAACCT	GGGGAGAAGA	GCACTGTCTT	CACCATCTTT	1800
	GGCTCGAACA	AAGAGAATGT	CCACAGCTGG	CTGATCCTCC	AGGTCAATGC	CACGGATGCC	1860
45	TTGGGAGTTC	CCTGCACAGA	GAATGACTAC	AAGCTGTGGT	CACCATCTGA	TGAGCGGGG	1920
	AATGAGTGTT	TGCTGGGACA	CAAGACTGTT	TTCAAACGGC	GGACCCCCCA	TGCCACATGC	1980
	TTCAATGGAG	AGGACTTTGA	CAGGCCGGTG	GTCGTGTCCA	ACTGCTCCTG	CACCCGGGAG	2040
50	GACTATGAGT	GTGACTTCGG	TTTCAAGATG	AGTGAAGATT	TGTCATTAGA	GGTTTGTGTT	2100
	CCAGATCCGG	AATTTTCTGG	AAAGTCATAC	TCCCCTCCTG	TGCCTTGCCC	TGTGGGTTCT	2160

	ACTTACAGGA	GAACGAGAGG	CTACCGGAAG	ATTTCTGGGG	ACACTTGTAG	CGGAGGAGAT	2220
	GTTGAAGCGC	GACTGGAAGG	AGAGCTGGTC	CCCTGTCCCC	TGGCAGAAGA	GAACGAGTTC	2280
5	ATTCTGTATG	CTGTGAGGAA	ATCCATCTAC	CGCTATGACC	TGGCCTCGGG	AGCCACCGAG	2340
	CAGTTGCCTC	TCACCGGGCT	ACGGGCAGCA	GTGGCCCTGG	ACTTTGACTA	TGAGCACAAC -	2400
10	TGTTTGTATT	GGTCCGACCT	GGCCTTGGAC	GTCATCCAGC	GCCTCTGTTT	GAATGGAAGC	2460
	ACAGGGCAAG	AGGTGATCAT	CAATTCTGGC	CTGGAGACAG	TAGAAGCTTT	GGCTTTTGAA	2520
	CCCCTCAGCC	AGCTGCTTTA	CTGGGTAGAT	GCAGGCTTCA	AAAAGATTGA	GGTAGCTAAT	2580
15	CCAGATGGCG	ACTTCCGACT	CACAATCGTC	AATTCCTCTG	TGCTTGATCG	TCCCAGGGCT	2640
	CTGGTCCTCG	TGCCCCAAGA	GGGGGTGATG	TTCTGGACAG	ACTGGGGAGA	CCTGAAGCCT	2700
	GGGATTTATC	GGAGCAATAT	GGATGGTTCT	GCTGCCTATC	ACCTGGTGTC	TGAGGATGTG	2760
20	AAGTGGCCCA	ATGGCATCTC	TGTGGACGAC	CAGTGGATTT	ACTGGACGGA	TGCCTACCTG	2820
	GAGTGCATAG	AGCGGATCAC	GTTCAGTGGC	CAGCAGCGCT	CTGTCATTCT	GGACAACCTC	2880
25	CCGCACCCCT	ATGCCATTGC	TGTCTTTAAG	AATGAAATCT	ACTGGGATGA	CTGGTCACAG	2940
	CTCAGCATAT	TCCGAGCTTC	CAAATACAGT	GGGTCCCAGA	TGGAGATTCT	GGCAAACCAG	3000
	CTCACGGGGC	TCATGGACAT	GAAGATTTTC	TACAAGGGGA	AGAACACTGG	AAGCAATGCC	3060
30	TGTGTGCCCA	GGCCATGCAG	CCTGCTGTGC	CTGCCCAAGG	CCAACAACAG	TAGAAGCTGC	3120
	AGGTGTCCAG	AGGATGTGTC	CAGCAGTGTG	CTTCCATCAG	GGGACCTGAT	GTGTGACTGC	3180
	CCTCAGGGCT	ATCAGCTCAA	GAACAATACC	TGTGTCAAAG	AAGAGAACAC	CTGTCTTCGC	3240
35	AACCAGTATC	GCTGCAGCAA	CGGGAACTGT	ATCAACAGCA	TTTGGTGGTG	TGACTTTGAC	3300
	AACGACTGTG	GAGACATGAG	CGATGAGAGA	AACTGCCCTA	CCACCATCTO	TGACCTGGAC	3360
40	ACCCAGTTTC	GTTGCCAGGA	GTCTGGGACT	TGTATCCCAC	TGTCCTATAA	ATGTGACCTT	3420
40	GAGGATGACT	GTGGAGACAA	CAGTGATGAA	AGTCATTGTG	AAATGCACCA	GTGCCGGAGT	3480
	GACGAGTACA	ACTGCAGTTC	CGGCATGTGC	ATCCGCTCCT	CCTGGGTATO	TGACGGGGAC	3540
45	AACGACTGCA	GGGACTGGTC	TGATGAAGC	CAACTGTACCG	CCATCTATCA	CACCTGTGAG	3600
	GCCTCCAACT	TCCAGTGCCC	AAACGGGCAC	TGCATCCCCC	AGCGGTGGG	GTGTGACGGG	3660
	GATACGGACT	GCCAGGATGC	TTCCGATGA(G GATCCAGTCA	ACTGTGAGA	GAAGTGCAAT	3720
50	GGATTCCGCT	GCCCAAACGC	CACTTGCAT	C CCATCÇAGCA	AACATTGTG/	TGGTCTGCGT	3780
	ርል ተተርርተርተር	: <u>ለጥርርር</u> ተርርር	ላ ፕሮልልሮልርርልር	TOCGAGCCC	TCTGTACGC	A CTTCATGGAC	3840

	TTTGTGTGTA	AGAACCGCCA	GCAGTGCCTG	TTCCACTCCA	TGGTCTGTGA	CGGATATCATC	3900
5	CAGTGCCGCG	ACGGGTCCGA	TGAGGATGCG	GCGTTTGCAG	GATGCTCCCA	AGATCCTGAG	3960
	TTCCACAAGG	TATGTGATGA	GTTCGGTTTC	CAGTGTCAGA	ATGGAGTGTG	CATCAGTTTG	4020
	ATTTGGAAGT	GCGACGGGAT	GGATGATTGC	GGCGATTATT	CTGATGAAGC	CAACTGCGAA	4080
10	AACCCCACAG	AAGCCCCAAA	CTGCTCCCGC	TACTTCCAGT	TTCGGTGTGA	GAATGGCCAC	4140
	TGCATCCCCA	ACAGATGGAA	ATGTGACAGG	GAGAACGACT	GTGGGGACTG	GTCTGATGAG	4200
	AAGGATTGTG	GAGATTCACA	TATTCTTCCC	TTCTCGACTC	CTGGGCCCTC	CACGTGTCTG	4260
15	CCCAATTACT	ACCGCTGCAG	CAGTGGGACC	TGCGTGATGG	ACACCTGGGT	GTGCGACGGG	4320
	TACCGAGATT	GTGCAGATGG	CTCTGACGAG	GAAGCCTGCC	CCTTGCTTGC	AAACGTCACT	4380
20	GCTGCCTCCA	CTCCCACCCA	ACTTGGGCGA	TGTGACCGAT	TTGAGTTCGA	ATGCCACCAA	4440
20	CCGAAGACGT	GTATTCCCAA	CTGGAAGCGC	TGTGACGGCC	ACCAAGATTG	CCAGGATGGC	4500
	CGGGACGAGG	CCAATTGCCC	CACACACAGC	ACCTTGACTT	GCATGAGCAG	GGAGTTCCAG	4560
25	TGCGAGGACG	GGGAGGCCTG	CATTGTGCTC	TCGGAGCGCT	GCGACGGCTT	CCTGGACTGC	4620
	TCGGACGAGA	GCGATGAAAA	GGCCTGCAGT	GATGAGTTGA	CTGTGTACAA	AGTACAGAAT	4680
	CTTCAGTGGA	CAGCTGACTT	CTCTGGGGAT	GTGACTTTGA	CCTGGATGAG	GCCCAAAAA	4740
30	ATGCCCTCTG	CATCTTGTGT	ATATAATGTC	TACTACAGGG	TGGTTGGAGA	GAGCATATGG	4800
	AAGACTCTGG	AGACCCACAG	CAATAAGACA	AACACTGTAT	TAAAAGTCTT	GAAACCAGAT	4860
	ACCACGTATC	AGGTTAAAGT	ACAGGTTCAG	TGTCTCAGCA	AGGCACACAA	CACCAATGAC	4920
35	TTTGTGACCC	TGAGGACCCC	AGAGGGATTG	CCAGATGCCC	CTCGAAATCT	CCAGCTGTCA	4980
	CTCCCCAGGG	AAGCAGAAGG	TGTGATTGTA	GGCCACTGGG	CTCCTCCCAT	CCACACCCAT	5040
40	GGCCTCATCC	GTGAGTACAT	TGTAGAATAC	AGCAGGAGTG	GTTCCAAGAT	GTGGGCCTCC	5100
	CAGAGGGCTG	CTAGTAACTT	TACAGAAATC	AAGAACTTAT	TGGTCAACAC	TCTATACACC	5160
	GTCAGAGTGG	CTGCGGTGAC	TAGTCGTGGA	ATAGGAAACT	GGAGCGATTC	TAAATCCATT	5220
45	ACCACCATAA	AAGGAAAAGT	GATCCCACCA	CCAGATATCC	ACATTGACAG	CTATGGTGAA	5280
	AATTATCTAA	GCTTCACCCT	GACCATGGAG	AGTGATATCA	AGGTGAATGG	CTATGTGGTG	5340
	AACCTTTTCT	GGGCATTTGA	CACCCACAAG	CAAGAGAGGA	GAACTTTGAA	CTTCCGAGGA	5400
50	AGCATATTGT	CACACAAAGT	TGGCAATCTG	ACAGCTCATA	CATCCTATGA	GATTTCTGCC	5460
	TGGGCCAAGA	CTGACTTGGG	GGATAGCCCT	CTGGCATTTG	AGCATGTTAT	GACCAGAGGG	5520

	GTTCGCCCAC CTGCACCTA	G CCTCAAGGCC	AAAGCCATCA	ACCAGACTGC	AGTGGAATGT	5580
,	ACCTGGACCG GCCCCCGGA	A TGTGGTTTAT	GGTATTTTCT	ATGCCACGTC	CTTTCTTGAC	5640
•	CTCTATCGCA ACCCGAAGA	G CTTGACTACT	TCACTCCACA	ACAAGACGGT	CATTGTCAGT	5700
	AAGGATGAGC AGTATTTG1	T TCTGGTCCGT	GTAGTGGTAC	CCTACCAGGG	GCCATCCTCT	5760
10	GACTACGTTG TAGTGAAGA	T GATCCCGGAC	AGCAGGCTTC	CACCCCGTCA	CCTGCATGTG	5820
	GTTCATACGG GCAAAACCT	C CGTGGTCATC	AAGTGGGAAT	CACCGTATGA	CTCTCCTGAC	5880
	CAGGACTTGT TGTATGCA	T TGCAGTCAAA	GATCTCATAA	GAAAGACTGA	CAGGAGCTAC	5940
15	AAAGTAAAAT CCCGTAACA	G CACTGTGGAA	TACACCCTTA	ACAAGTTGGA	GCCTGGCGGG	6000
	AAATACCACA TCATTGTCC	A ACTGGGGAAC	ATGAGCAAAG	ATTCCAGCAT	AAAAATTACC	6060
00	ACAGTTTCAT TATCAGCAG	C TGATGCCTTA	AAAATCATAA	CAGAAAATGA	TCATGTTCTT	6120
20	CTGTTTTGGA AAAGCCTGG	C TTTAAAGGAA	AAGCATTTTA	ATGAAAGCAG	GGGCTATGAG	6180
	ATACACATGT TTGATAGT	C CATGAATATC	ACAGCTTACC	TTGGGAATAC	TACTGACAAT	6240
25	TTCTTTAAAA TTTCCAACO	T GAAGATGGGT	CATAATTACA	CGTTCACCGT	CCAAGCAAGA	6300
,	TGCCTTTTTG GCAACCAGA	T CTGTGGGGAG	CCTGCCATCC	TGCTGTACGA	TGAGCTGGGG	6360
	TCTCGTGCAG ATGCATCTC	C AACGCAGGCT	GCCAGATCTA	CGGATGTTGC	TGCTGTGGTG	6420
30	GTGCCCATCT TATTCCTGA	T ACTGCTGAGC	CTGGGGGTGG	GGTTTGCCAT	CCTGTACACG	6480
	AAGCACCGGA GGCTGCAGA	G CAGCTTCACC	GCCTTCGCCA	ACAGCCACTA	CAGCTCCAGG	6540
	CTGGGGTCCG CAATCTTCT	C CTCTGGGGAT	GACCTGGGGG	AAGATGATGA	AGATGCCCCT	6600
35	ATGATAACTG GATTTTCAC	A TGACGTCCCC	ATGGTGATAG	CC		6642
	Sequence ID No.	5				
40	Length of the Se	quence: 22	14			
	Type: amino acid					
	Topology: linear					
45	Molecular type:	Protein				
	Sequence:					
	Met Ala Thr Arg Ser	Ser Arg Arg	Glu Ser Arg	g Leu Pro Pl	ne Leu Phe	
50	5	•	10		15	
	Thr Leu Val Ala Leu	Leu Pro Pro	Gly Ala Lei	u Cys Glu V	al Trp Thr	

				20)				25	5				30)	
5	Gli	n Arg	Leu	His	Gly	Gly	/ Ser	Ala	Pro) Leu	Pro	Gln	Asp	Arg	Gly	Phe
			35					40					45			
	Lei	ı Val	Val	Gln	Gly	Asp	Pro	Arg	Glu	Leu	Arg	Leu	Тгр	Ala	Arg	Gly
10		50					55					60				
	Asp	Ala	Arg	Gly	Ala	Ser	Arg	Ala	Asp	Glu	Lys	Pro	Leu	Arg	Arg	Lys
15	65					70					75					80
,,	Arg	Ser	Ala	Ala	Leu	Gln	Pro	Glu	Pro	lle	Lys	Val	Туг	Gly	Gln	Val
					85					90					95	
20	Ser	Leu	Asn	Asp	Ser	His	Asn	Gln	Met	Val	Val	His	Trp	Ala	Gly	Glu
				100					105					110		
	Lys	Ser	Asn	Val	Ile	Val	Ala	Leu	Ala	Arg	Asp	Ser	Leu	Ala	Leu	Ala
25			115					120					125			
	Arg	Pro	Lys	Ser	Ser	Asp	Val	Туг	Val	Ser	Tyr	Asp	Туг	Gly	Lys	Ser
		130					135					140				
30	Phe	Lys	Lys	Ιlе	Ser	Asp	Lys	Leu	Asn	Phe	Gly	Leu	Gly	Asn	Arg	Ser
	145					150					155					160
35	Glu	Ala	Val	Ile	Ala	Gln	Phe	Туг	His	Ser	Pro	Ala	Asp	Asn	Lys	Arg
33					165					170					175	
	Tyr	He	Phe	Ala	Asp	Ala	Туг	Ala	Gln	Tyr	Leu	Trp	Πle	Thr	Phe	Asp
40				180					185					190		
	Phe	Cys	Asn	Thr	Leu	Gln	Gly	Phe	Ser	Ιle	Pro	Phe	Arg	Ala	Ala	Asp
			195					200					205			
45	Leu	Leu	Leu	His	Ser	Lys	Ala	Ser	Asn	Leu	Leu	Leu	Gly	Phe	Asp	Arg
		210					215					220				
	Ser	His	Pro	Asn	Lys	Gln	Leu	Trp	Lys	Ser	Asp	Asp	Phe	Gly	Gln	Thr
50	225					230					235					240
	Trp	He	Met	lle	Gln	Glu	His	Val	Lys	Ser	Phe	Ser	Тгр	Gly	He	Asp

					245					250					255	
5	Pro	Tyr	Asp	Lys	Pro	Asn	Thr	He	Туг	He	Glu	Arg	His	Glu	Pro	Ser
				260		•			265					270		
	Gly	Tyr	Ser	Thr	Val	Phe	Arg	Ser	Thr	Asp	Phe	Phe	Gln	Ser	Arg	Glu
10			275					280					285			
	Asn	Gln	Glu	Val	Пе	Leu	Glu	Glu	Val	Arg	Asp	Phe	Gln	Leu	Arg	Asp
		290					295					300				
15	Lys	Туг	Met	Phe	Ala	Thr	Lys	Val	Val	His	Leu	Leu	Gly	Ser	Glu	Gln
	305					310					315					320
	Gln	Ser	Ser	Val	Gln	Leu	Trp	Val	Ser	Phe	Gly	Arg	Lys	Pro	Met	Arg
20					325					330					335	
	Ala	Ala	Gln	Phe	Val	Thr	Arg	His	Pro	He	Asn	Glu	Tyr	Туг	Ile	Ala
25				340					345					350		
	Asp	Ala	Ser	Glu	Asp	Gln	Val	Phe	Val	Cys	Val	Ser	His	Ser	Asn	Asn
	•		355					360					365			
30	Arg	Thr		Leu	Туг	He	Ser		Ala	Glu	Gly	Leu	Lys	Phe	Ser	Leu
	_	370					375					380				
	Ser	Leu	Glu	Asn	Val	Leu		Tyr	Ser	Pro	Gly	Gly	Ala	Gly	Ser	Asp
35	385					390					395					400
		Leu	Val	Arg	Туг		Ala	Asn	Glu	Pro	Phe	Ala	Asp	Phe	His	Arg
				J	405					410					415	
40	Val	Glu	Glv	Leu			Val	Туг	ile	Ala	Thr	Leu	lle	Asn	Gly	Ser
			,	420				•	425					430		
45	Met	Asn	Glu		Asn	Met	Arg	Ser			Thr	Phe	Asp	Lys	Gly	Gly
,5			435					440					445			
	Thr	Trp			Leu	Gln	Ala			Phe	Thr	Gly			Glu	Lys
50		450					455					460		-		
	[]e	Asn		Glu	Leu	Ser			Cys	Ser	Leu			Ala	Gln	Arg
			,0	~ 			••		•							•

	465					470					475					480
5	Leu	Ser	Gln	Leu	Leu	Asn	Leu	Gln	Leu	Arg	Arg	Met	Pro	Ile	Leu	Ser
					485					490					495	
	Lys	Glu	Ser	Ala	Pro	Gly	Leu	Пе	Ile	Ala	Thr	Gly	Ser	Val	Gly	Lys
10				500					505					510		
	Asn	Leu	Ala	Ser	Lys	Thr	Asn	Val	Tyr	Ile	Ser	Ser	Ser	Ala	Gly	Ala
			515					520					525			
15	Arg	Trp	Arg	Glu	Ala	Leu	Pro	Gly	Pro	His	Tyr	Tyr	Thr	Trp	Gly	Asp
		530					535					540				
	His	Gly	Gly	Ιlе	lle	Thr	Ala	lle	Ala	Gln	Gly	Met	Glu	Thr	Asn	Glu
20	545					550					555					560
	Leu	Lys	Tyr	Ser	Thr	Asn	Glu	Gly	Glu	Thr	Trp	Lys	Thr	Phe	Пe	Phe
25					565					570					575	
	Ser	Glu	Lys	Pro	Val	Phe	Val	Tyr	Gly	Leu	Leu	Thr	Glu	Pro	Gly	Glu
				580					585				-	590		
30	Lys	Ser	Thr	Val	Phe	Thr	He	Phe	Gly	Ser	Asn	Lys	Glu	Asn	Val	His
			595					600					605			
	Ser	Trp	Leu	He	Leu	Gln	Val	Asn	Ala	Thr	Asp	Ala	Leu	Gly	Val	Pro
35		610					615					620				
	Cys	Thr	Glu	Asn	Asp	Туг	Lys	Leu	Trp	Ser	Pro	Ser	Asp	Glu	Arg	Gly
40	625					630					635					640
40	Asn	Glu	Cys	Leu	Leu	Gly	His	Lys	Thr	Val	Phe	Lys	Arg	Arg	Thr	Pro
					645					650					655	
4 5	His	Ala	Thr	Cys	Phe	Asn	Gly	Glu	Asp	Phe	Asp	Arg	Pro	Val	Val	Val
				660					665					670		
	Ser	Asn	Cys	Ser	Cys	Thr	Arg	Glu	Asp	Tyr	Glu	Cys	Asp	Phe	Gly	Phe
50			675					680					685			
	Lys	Met	Ser	Glu	Asp	Leu	Ser	Leu	Glu	Val	Cys	Val	Pro	Asp	Pro	Glu

		690					695					700					
_	Phe	Ser	Gly	Lys	Ser	Туг	Ser	Pro	Pro	Val	Pro	Cys	Pro	Val	Gly	Se	r
5	705					710					715					72	0
	Thr	Туг	Arg	Arg	Thr	Arg	Gly	Tyr	Arg	Lys	Ile	Ser	Gly	Asp	Thr	Су	s
10					725					730					735		
	Ser	Gly	Gly	Asp	Val	Glu	Ala	Arg	Leu	Glu	Gly	Glu	Leu	Val	Pro	Су	S
				740					745					7 50			
15	Pro	Leu	Ala	Glu	Glu	Asn	Glu	Phe	He	Leu	Tyr	Ala	Val	Arg	Lys	Se	r
			755					760					765				
	Ile	Туг	Arg	Туг	Asp	Leu	Ala	Ser	Gly	Ala	Thr	Glu	Gln	Leu	Pro	Le	eu
20		770					775					780					
	Thr	Gly	Leu	Arg	Ala	Ala	Val	Ala	Leu	Asp	Phe	Asp	Tyr	Glu	His	s As	sn
25	785					790					795	,				8	00
25	Cys	Leu	Туг	Trp	Ser	Asp	Leu	Ala	Leu	ı Asp	Val	Ιle	Glr	n Arg	g Lei	1 C	ys
					805	5				810)				819	ō	
30	Leu	Asn	Gly	/ Ser	Thr	Gly	Glr	Glu	ı Val	H	e Ile	e Ası	n Sei	r Gl	y Le	цG	lu
				820)				825	ō				83)		
	Thr	· Val	Gle	ı Ala	a Lei	ı Ala	a Pho	e Gli	ı Pro	o Le	u Se	r Gli	n Le	u Le	u Ty	r T	rp
35			83	5				84	0				84	5			
	۷a	l As	p Al	a Gl	y Ph	e Ly	s Ly	s II	e Gl	u Va	l Al	a As	n Pr	o As	p Gl	y l	Asp
		85	0				85	5				86	0				
40	Ph	e Ar	g Le	u Th	r Il	e Va	l As	n Se	r Se	er Va	al Le	eu As	sp Ar	g Pi	o A	g	Ala
·	86	5				87	0'				8	75					880
45	Le	eu Va	al Le	eu Va	al Pr	o Gi	n G	u G1	ly Va	al Me	et Pi	ne Ti	rp Tl	hr As	sр Т	rp	Gly
45						35					90					95	
	As	sp Le	eu Ly	ys P	ro G	ly I	le T	yr Ai	rg S	er A	sn M	et A	sp G	ly S	er A	la	Ala
50	-				00					05					10		
	T	yr H	is L	eu V	al S	er G	lu A	sp V	al L	ys T	rp P	ro A	sn G	ly I	le S	er	Val

			915	5				92)				92	5		
5	Asp	Asp	Glr	1rp	ı Ile	∈ Ту	r Tr	p Thi	As	p Al	а Ту	r Le	u Gl	u Cy:	s II	e Glu
-		930					93					94				
	Arg	Ile	Thr	Phe	: Ser	Gly	y Gli	n Glr	n Arg	g Sei	r Val	11	e Le	u Ası	As	n Leu
10	945					950					955					960
	Pro	His	Pro	Tyr	Ala	Пε	e Ala	a Val	Phe	e Lys	s Asr	Gli	u Ile	e Tyr	Tr	g Asp
					965					970					975	
15	Asp	Trp	Ser	Gln	Leu	Ser	He	Phe	: Arg	Ala	a Ser	Lys	s Tyı	Ser	Gly	/ Ser
				980					985					990		
	Gln	Met	Glu	lle	Leu	Ala	Asn	Gln	Leu	Thr	Gly	Lei	ı Met	: Asp	Met	: Lys
20			995					100					100			
	Ile	Phe	Tyr	Lys	Gly	Lys	Asn	Thr	Gly	Ser	Asn	Ala	a Cys	. Val	Pro	Arg
25		1010					101					102				
	Pro	Cys	Ser	Leu	Leu	Cys	Leu	Pro	Lys	Ala	Asn	Asr	ı Ser	Arg	Ser	Cys
	1025					103					103					1040
30	Arg	Cys	Pro	Glu	Asp	Val	Ser	Ser	Ser	Val	Leu	Pro	Ser	Gly	Asp	Leu
					1049					105					105	
	Met	Cys	Asp	Cys	Pro	Gln	Gly	Туг	Gln	Leu	Lys	Asn	Asn	Thr	Cys	Val
35				1060					106					107		
	Lys (Glu (Glu .	Asn '	Thr	Cys	Leu	Arg	Asn	Gln	Туг	Arg	Cys	Ser	Asn	Gly
40			1075					1080					1085			
	Asn C	ys i	lle <i>i</i>	Asn S	Ser	lle	Тгр	Trp	Cys	Asp	Phe	Asp	Asn	Asp	Cys	Gly
	1	090					1095					1100)			
45	Asp M	let S	Ser A	Asp (Glu <i>i</i>	Arg.	Asn	Cys	Pro	Thr	Thr	lle	Cys	Asp	Leu	Asp
	1105					1110					1115					1120
			he A	rg C			Glu .	Ser (Gly '				Pro	Leu	Ser	
50	1105		he A				Glu .	Ser (Pro		Ser 1135	Tyr
50	1105	ln F		1	Cys (Gln (Thr 1130	Cys	lle			1135	Туг

				1140)				1145	5				1150)		
	Cys	Glu	Met	His	G1 n	Cys	Arg	Ser	Asp	Glu	Tyr	Asn	Cys	Ser	Ser	G	у
5			1155	5				1160)				1165	5			
	Met	Cys	He	Arg	Ser	Ser	Trp	Val	Cys	Asp	Gly	Asp	Asn	Asp	Cys	A	rg
10		117	0				1175	5				1180)				
	Asp	Тгр	Ser	Asp	Glu	Ala	Asn	Cys	Thr	Ala	He	Tyr	His	Thr	Cys	G	lu
	118	5				119	0				119	5				1	200
15	Ala	Ser	Asn	Phe	Gln	Cys	Arg	Asn	Gly	His	Cys	Пe	Pro	Gln	Arg	Ţ	rp
					120	5				121	0				121	15	
	Ala	Cys	Asp	Gly	Asp	Thr	Asp	Cys	Gln	Asp	Gly	Ser	Asp	Glu	Ası	P	ro
20				122	0.				122	5				123	0		
	Val	Asn	Cys	Glu	Lys	Lys	Cys	Asn	Gly	Phe	Arg	Cys	Pro	Asr	Gl	у 1	`hr
25			123	35				124	0.				124	15			
	Cys	: Ile	e Pro	Ser	Ser	Lys	His	Cys	Asp	Gly	Leu	Arg	g Asp	Cys	s Se	r A	lsp
		125	50				125	55				126	60				
30	Gly	/ Se	r Ası	o Glu	ı Glı	n His	s Cys	s Glu	ı Pro	o Lei	ı Cys	s Thi	r His	s Ph	e Me	t /	Asp
	120	35				12'	70				12	75					1280
	Ph	e Va	l Cy	s Ly:	s Asi	n Ar	g Gl	n Gli	n Cy	s Lei	ı Pho	e Hi	s Se	r Me			
35					12					12						95	
	As	p Gl	y Il	e II	e Gl	n Cy	s Ar	g As	p Gl	y Se	r As	p Gl	u As	p Al	a Al	a	Phe
40				. 13	300				13	305				13	310		
	Al	a Gl	y Cy	ıs Se	er Gl	n As	sp Pr	o Gl	u Pł	ne Hi	s Ly	's Va	ıl Cy	s As	sp G	l u	Phe
			13	315				13	320				13	325			
45	G1	y Ph	ne Gl	n Cy	s Gl	n As	sn Gl	y Va	ıl Cy	/s Il	e Se	er Le	eu II	e Ti	rp L	ys	Cys
			330					335					340				
	As	sp G	ly Me	et As	sp As	sp Cy	/s Gl	y As	зр Ту	yr Se			lu Al	la As	sn C	ys	
50		345					350					355					1360
	As	sn Pi	ro Ti	hr G	lu Al	la Pi	ro As	sn Cy	ys S	er Ai	g Ty	ır Pl	he G	ln P	he A	rg	Cys

		1365	1370)	1375
_	Glu Asn Gly	His Cys Ile	Pro Asn Arg Trp	Lys Cys Asp A	
5		1380	1385		390
	Asp Cys Gly	Asp Trp Ser	Asp Glu Lys Asp	Cys Gly Asp Se	er His Ile
10	1399		1400	1405	
	Leu Pro Phe	Ser Thr Pro	Gly Pro Ser Thr	Cys Leu Pro As	sn Tyr Tyr
	1410		1415	1420	•
15	Arg Cys Ser	Ser Gly Thr	Cys Val Met Asp	Thr Trp Val Cy	's Asp Gly
	1425	1430		1435	1440
	Tyr Arg Asp	Cys Ala Asp	Gly Ser Asp Glu (Glu Ala Cys Pr	
20		1445	1450		1455
	Ala Asn Val	Thr Ala Ala	Ser Thr Pro Thr (in Leu Gly Ar	
<i>25</i>		1460	1465	14'	
	Arg Phe Glu	Phe Glu Cys I	His Gln Pro Lys T	hr Cys Ile Pr	o Asn Trp
	1475		1480	1485	•
30	Lys Arg Cys /	Asp Gly His (Gln Asp Cys Gln A	sp Gly Arg Asj	o Glu Ala
	1490		1495	1500	
	Asn Cys Pro 1	Chr His Ser 1	Thr Leu Thr Cys M	et Ser Arg Glı	ı Phe Gln
35	1505	1510		515	1520
	Cys Glu Asp (Sly Glu Ala C	Cys Ile Val Leu S	er Glu Arg Cys	
40		1525	1530		1535
	Phe Leu Asp C	ys Ser Asp G	lu Ser Asp Glu Ly	s Ala Cys Ser	Asp Glu
		540	1545	155	
45	Leu Thr Val T	yr Lys Val G	ln Asn Leu Gln Tr		
	1555		1560	1565	
	Gly Asp Val T	hr Leu Thr Ti	rp Met Arg Pro Ly	s Lys Met Pro	Ser Ala
50	1570		575	1580	
	Ser Cys Val Ty	r Asn Val Ty	yr Tyr Arg Val Va	l Gly Glu Ser	Ile Trp
55					-

	1589	5				1590)				1595	5		٠		1600
5	Lys	Thr	Leu	Glu	Thr	His	Ser	Asn	Lys	Thr	Asn	îhr	Val	Leu	Lys	Val
5					1605	5				1610)				1615	5
	Leu	Lys	Pro	Asp	Thr	Thr	Туг	Gln	Val	Lys	Val	G1 n	Val	Gln	Cys	Leu
10				1620)				1629	5				1630)	
	Ser	Lys	Ala	His	Asn	Thr	Asn	Asp	Phe	Val	Thr	Leu	Arg	Thr	Pro	Glu
			1635	5				1640)				1645	5		
15	Gly	Leu	Pro	Asp	Ala	Рго	Arg	Asn	Leu	Gln	Leu	Ser	Leu	Pro	Arg	Glu
		1650)				1655	5				1660)			
	Ala	Glu	Gly	Val	Ile	Val	Gly	His	Trp	Ala	Pro	Pro	He	His	Thr	His
20	1669	5				1670)	-			1675	5				1680
	Gly	Leu	Ιle	Arg	Glu	Tyr	Ιle	Val	Glu	Tyr	Ser	Arg	Ser	Gly	Ser	Lys
25					1689	5				1690)				1695	5
25	Met	Trp	Ala	Ser	Gln	Arg	Ala	Ala	Ser	Asn	Phe	Thr	Glu	Ile	Lys	Asn
				1700)				1709	5				1710)	
30	Leu	Leu	Val			Leu	Tyr	Thr			Val	Ala	Ala			Ser
30	Leu	Leu	Val	Asn		Leu	Tyr	Thr 1720	Val		Val	Ala	Ala 1729	Val		Ser
30	•	Leu Gly	1715	Asn 5	Thr			1720	Val)	Arg			172	Val 5	Thr	
35	•		1715 Ile	Asn 5	Thr			1720 Asp	Val)	Arg			1725 Thr	Val 5	Thr	
	Arg	Gly	1715 Ile)	Asn 5 Gly	Thr Asn	Trp	Ser 1735	172(Asp	Val) Ser	Arg Lys	Ser	11e	1729 Thr	Val 5 Thr	Thr	Lys
	Arg	Gly 1730 Lys	1715 Ile)	Asn 5 Gly	Thr Asn	Trp	Ser 1735 Pro	172(Asp	Val) Ser	Arg Lys	Ser	Ile 1740 Asp	1729 Thr	Val 5 Thr	Thr	Lys
	Arg Gly	Gly 1730 Lys	1715 	Asn Gly Ile	Thr Asn Pro	Trp Pro 1750	Ser 1735 Pro	1720 Asp 5 Asp	Val) Ser Ile	Arg Lys His	Ser 11e 175	11e 1740 Asp	1729 Thr O Ser	Val Thr Tyr	Thr lle Gly	Lys Glu 1760
35	Arg Gly	Gly 1730 Lys	1715 	Asn Gly Ile	Thr Asn Pro	Trp Pro 1750 Thr	Ser 1735 Pro	1720 Asp 5 Asp	Val) Ser Ile	Arg Lys His	Ser He 1759 Ser	11e 1740 Asp	1729 Thr O Ser	Val Thr Tyr	Thr lle Gly	Lys Glu 1760 Asn
35	Gly 1745 Asn	Gly 1730 Lys 5 Tyr	1715 Ile Val	Asn Gly Ile Ser	Thr Asn Pro Phe 176	Trp Pro 1750 Thr	Ser 1735 Pro) Leu	1720 Asp S Asp Thr	Val) Ser Ile Met	Arg Lys His Glu 177	Ser He 1759 Ser	Ile 1740 Asp 5 Asp	Thr O Ser	Val Thr Tyr Lys	Thr lle Gly Val	Lys Glu 1760 Asn
35	Gly 1745 Asn	Gly 1730 Lys 5 Tyr	1715 Ile Val	Asn Gly Ile Ser	Thr Asn Pro Phe 1763	Trp Pro 1750 Thr	Ser 1735 Pro) Leu	1720 Asp S Asp Thr	Val) Ser Ile Met	Arg Lys His Glu 177	Ser He 1759 Ser	Ile 1740 Asp 5 Asp	Thr O Ser	Val Thr Tyr Lys	Thr Ile Gly Val 177 Gln	Lys Glu 1760 Asn 5
35	Gly 1745 Asn	Gly 1730 Lys 5 Tyr	1715 Ile) Val Leu	Asn Gly Ile Ser Val	Thr Asn Pro Phe 1766 Asn	Trp Pro 1750 Thr 5	Ser 1735 Pro) Leu Phe	1720 Asp 5 Asp Thr	Val) Ser Ile Met Ala	Arg Lys His Glu 1770 Phe	Ile 1755 Ser O	Ile 1744 Asp 5 Asp	Thr O Ser Ile	Val Thr Tyr Lys Lys 179	Thr Ile Gly Val 177 Gln 0	Lys Glu 1760 Asn 5 Glu
35	Gly 1745 Asn	Gly 1730 Lys Tyr	1715 Ile) Val Leu	Asn Gly Ile Ser Val 1780 Leu	Thr Asn Pro Phe 1766 Asn	Trp Pro 1750 Thr 5	Ser 1735 Pro) Leu Phe	1720 Asp 5 Asp Thr	Val) Ser Ile Met Ala 178 Ser	Arg Lys His Glu 1770 Phe	Ile 1755 Ser O	Ile 1744 Asp 5 Asp	Thr O Ser Ile	Val Thr Tyr Lys Lys 179 Lys	Thr Ile Gly Val 177 Gln 0	Lys Glu 1760 Asn 5 Glu

		18	10				18	15				18	20			
5	Ası	Le	u Gl	y As	p Se	r Pro	b Le	u Al	a Ph	e Glu	ı His	s Va	l Me	t Th	r Arį	gGly
	182					183					183					1840
	Val	Ar	g Pro	o Pr	o Ala	a Pro	Se	r Le	ц Ly:	s Ala	a Lys	. Ala	a Ile	e Ası	n Glr	Thr
10					184					185					185	
	Ala	\ Va	l Glu	ı Cy:	s Thr	Trp	Th	r Gl	у Рг	o Arg	, Asn	Val	l Val	Tyr	Gly	lle
				186					186					187		
15	Phe	Туг	Ala	1 Thi	r Ser	Phe	: Lei	ı Ası	Lei	ı Tyr	Arg	Asr	Pro	Lys	Ser	Leu
			187					188					188			
20	Thr	Thr	Ser	Lei	ı His	Asn	Lys	Thr	Val	lle	Val	Ser	Lys	Asp	Glu	Gln
20		189					189					190				
	Tyr	Leu	Phe	Leu	ı Val	Arg	Val	Val	Val	Pro	Туг	Gln	Gly	Рго	Ser	Ser
25	1909					191					191					1920
	Asp	Tyr	Val	Val	Val	Lys	Met	Ιle	Pro	Asp	Ser	Arg	Leu	Рго	Pro	Arg
					192					193					193	
30	His	Leu	His	Val	Val	His	Thr	Gly	Lys	Thr	Ser	Va!	Val	Ile	Lys	Тгр
				194					194					195		
	Glu	Ser	Pro	Tyr	Asp	Ser	Pro	Asp	Gln	Asp	Leu	Leu	Туг	Ala	lle	Ala
35			195					196			. •		1969			
	Val	Lys	Asp	Leu	Ile	Arg	Lys	Thr	Asp	Arg	Ser	Туг	Lys	Val	Lys	Ser
10		1970					1979					1980				
	Arg	Asn	Ser	Thr	Val	Glu	Tyr	Thr	Leu	Asn	Lys	Leu	Glu	Pro	Gly	Gly
	1985					1990					1995					2000
15	Lys	Туг	His	He	Ιle	Val	Gln	Leu	Gly	Asn	Met	Ser	Lys	Asp	Ser	
					2005					2010			•		2015	
	[le	Lys	Ile	Thr	Thr	Val	Ser	Leu	Ser	Ala	Pro .	Asp	Ala			
ro .				2020					2025			-		2030		
	lle 1	ľhr	Glu	Asn	Asp i	His '	Val	Leu	Leu	Phe '	Trp i	Lys .				Leu
											•	• -				

			203	5				204	0				204	5		
_	Lys	Glu	Lys	His	Phe	Asn	Glu	Ser	Arg	Gly	Туг	Glu	Пe	His	Met	Phe
5		205	0				205	5				2060)			
	Asp	Ser	Ala	Met	Asn	Ile	Thr	Ala	Tyr	Leu	Gly	Asn	Thr	Thr	Asp	Asn
10	206	5				2070)				2075	5				2080
	Phe	Phe	Lys	Ile	Ser	Asn	Leu	Lys	Met	Gly	His	Asn	Tyr	Thr	Phe	Thr
					2085	5				2090)				2099	5
15	Val	Gln	Ala	Arg	Cys	Leu	Phe	Gly	Asn	Gln	Ile	Cys	Gly	Glu	Pro	Ala
				2100)				2109	5				2110)	
	Ile	Leu	Leu	Tyr	Asp	Glu	Leu	Gly	Ser	Gly	Ala	Asp	Ala	Ser	Ala	Thr
20			2115	5				2120)				2125	5		
	Gln	Ala	Ala	Arg	Ser	Thr	Asp	Val	Ala	Ala	Val	Val	Val	Pro	Ile	Leu
25		2130)				2135	5				2140)			
	Phe	Leu	lle	Leu	Leu	Ser	Leu	Gly	Val	Gly	Phe	Ala	Ile	Leu	Туг	Thr
	2145	5				2150	1				2155	5				2160
	214	,				2100	,				2100	•				2100
30			Arg	Arg	Leu			Ser	Phe	Thr			Ala	Asn	Ser	
30			Arg	Arg	Leu 2165	Gln		Ser	Phe	Thr 2170	Ala		Ala	Asn	Ser 217	His
30	Lys	His			2165	Gln 5	Ser		Phe	2170	Ala)	Phe			2179	His 5
35	Lys	His			2165 Leu	Gln 5	Ser			2170 Phe	Ala)	Phe			2179 Asp	His 5
	Lys Tyr	His Ser	Ser	Arg 2180	2165 Leu)	Gln Gly	Ser Ser	Ala	Ile	2170 Phe	Ala) Ser	Phe Ser	Gly	Asp 2190	2179 Asp	His 5 Leu
. 35	Lys Tyr	His Ser	Ser	Arg 2180 Asp	2165 Leu)	Gln Gly	Ser Ser	Ala	Ile 2185 Met	2170 Phe	Ala) Ser	Phe Ser	Gly	Asp 2190 Ser	2179 Asp	His 5 Leu
	Lys Tyr Gly	His Ser	Ser Asp 2195	Arg 2180 Asp	2165 Leu) Glu	Gln Gly Asp	Ser Ser	Ala Pro	Ile 2185 Met	2170 Phe	Ala) Ser	Phe Ser	Gly Phe	Asp 2190 Ser	2179 Asp	His 5 Leu
. 35	Lys Tyr Gly	His Ser Glu	Ser Asp 2195 Met	Arg 2180 Asp	2165 Leu) Glu	Gln Gly Asp	Ser Ser	Ala Pro	Ile 2185 Met	2170 Phe	Ala) Ser	Phe Ser	Gly Phe	Asp 2190 Ser	2179 Asp	His 5 Leu
. 35	Lys Tyr Gly	His Ser Glu Pro 2210	Ser Asp 2195 Met	Arg 2180 Asp Val	2165 Leu) Glu Ile	Gln Gly Asp	Ser Ser	Ala Pro	Ile 2185 Met	2170 Phe	Ala) Ser	Phe Ser	Gly Phe	Asp 2190 Ser	2179 Asp	His 5 Leu
35	Lys Tyr Gly Val	His Ser Glu Pro 2210	Ser Asp 2195 Met	Arg 2180 Asp Val	2165 Leu) Glu Ile	Gln Gly Asp Ala	Ser Ser Ala	Ala Pro 2200	Ile 2185 Met)	2170 Phe	Ala) Ser	Phe Ser	Gly Phe	Asp 2190 Ser	2179 Asp	His 5 Leu
35	Lys Tyr Gly Val	His Ser Glu Pro 2210 uenc gth	Ser Asp 2195 Met) ce I	Arg 2180 Asp Val Val	2165 Leu) Glu Ile Seq	Gln Gly Asp Ala	Ser Ser Ala	Ala Pro 2200	Ile 2185 Met)	2170 Phe	Ala) Ser	Phe Ser	Gly Phe	Asp 2190 Ser	2179 Asp	His 5 Leu
35	Lys Tyr Gly Val Seque	His Ser Glu Pro 2210 uenc gth	Ser Asp 2195 Met) ce I	Arg 2180 Asp Val D No the	2165 Leu) Glu Ile Seque	Gln Gly Asp Ala Juen	Ser Ser Ala	Ala Pro 2200	Ile 2185 Met)	2170 Phe	Ala) Ser	Phe Ser	Gly Phe	Asp 2190 Ser	2179 Asp	His 5 Leu

	Molecu	lar 1	type: d	DNA	to mRN	IA								
-	Featur	e:												
•	Name/	Key:	sig pe	ptid	е									
	Locat	ion:	8116	54										
10	Ident	ifica	ation m	etho	d: s									
	Name/	Key:	mat pe	ptide	е									
	Locat	ion:	1656	722										
15	Ident	ifica	ation m	etho	d: S									
•	Sequen	ce:												
								CCG	GCCC	AGCG	GC T	CTCC	TGGCC	23
20	TCGCGCT	rgca c	ATTCTCT	CC TGG	CGGCGG	C GCC	ACCT	GCA	GTAC	CGTT	°CG C	CCGA	ACATG	83
													Met	
25													1	
	GCG AC	CGG .	AGC AGC	AGG A	.GG GAG	TCG	CGA	CTC	CCG	TTC	CTA	TTC	ACC-	131
	Ala Thi	Arg	Ser Ser	Arg A	rg Glu	Ser	Arg	Leu	Pro	Phe	Leu	Phe	Thr	
30			5			10					15		٠	
	CTG GT	C GCA	CTG CTG	CCG C	CC GGA	GCT	CTC	TGC	GAA	GTC	TGG	ACG	CAG	179
	Leu Val	l Ala	Leu Leu	Pro P	ro Gly	Ala	Leu	Cys	Glu	Val	Тгр	Thr	Gln	
35		20			25					30				
	AGG CT	G CAC	GGC GGC	AGC G	GCG CCC	TTG	CCC	CAG	GAC	CGG	GGC	TTC	CTC	227
4 0	Arg Le	a His	Gly Gly	Ser A	la Pro	Leu	Pro	Gln	Asp	Arg	Gly	Phe	Leu	
	3	5			40				45					
	GTG GT	G CAG	GGC GAC	CCG C	GC GAG	CTG	CGG	CTG	TGG	GCG	CGC	GGG	GAT	275
45	Val Va	l Gln	Gly Asp	Pro A	rg Glu	Leu	Arg	Leu	Trp	Ala	Arg	Gly	Asp	
	50			55				60					65	
			GCG AGC											323
50	Ala Arı	gGly	Ala Ser	Arg A	la Asp	Glu	Lys	Pro	Leu	Arg	Arg	Lys	Arg	
			70				75					80		

	AGC	GCT	, GCC	CTC	CAC	ccc	GAG	CC	CAT	C AA	G GT	G TAC	GGA	CAG	GTT	AGT	371
5	Ser	Ala	Ala	Leu	Glr	n Pro	Glu	Pro	o II	e Ly:	s Va	l Tyr	Gly	Gln	Val	Ser	
				85	5				90	0				95			
	CTG	AAT	GAT	TCC	CAC	СААТ	CAG	AT(G GT	G GT	G CA	C TGG	GCT	GGA	GAG	AAA	419
10	Leu	Asn	Asp	Ser	His	Asn	Gln	Met	t Val	l Va	l His	s Trp	Ala	Gly	Ġlu	Lys	
			100					105	5				110				
	AGC	AAC	GTG	ATC	GTG	GCC	TTG	GCC	C CG/	A GAT	r AG(C CTG	GCA	TTG	GCG	AGG	467
15	Ser	Asn	Val	Пe	Val	Ala	Leu	Ala	ı Arş	g Asp	Ser	Leu	Ala	Leu	Ala	Arg	
		115					120					125					
20	CCC	AAG	AGC	AGT	GAT	GTG	TAC	GTC	тст	TAC	C GAC	ТАТ	GGA	AAA	TCA	TTC	515
20	Pro	Lys	Ser	Ser	Asp	Val	Tyr	Val	Ser	Tyr	Asp	Туг	Gly	Lys	Ser	Phe	
	130					135					140)				145	
25	AAG	AAA	ATT	TCA	GAC	AAG	TTA	AAC	777	r GGC	TTC	G GGA	AAT	AGG	AGT	GAA	563
	Lys	Lys	lle	Ser	Asp	Lys	Leu	Asn	Phe	e Gly	Leu	ıGly	Asn	Arg	Ser	Glu	
					150					155	5				160		
30	GCT	GTT	ATC	GCC	CAG	TTC	TAC	CAC	AGC	CCT	GC0	GAC	AAC	AAG	CGG	TAC	611
	Ala	Val	He	Ala	Gla	Phe	Tyr	His	Ser	Pro	Ala	Asp	Asn	Lys	Arg	Tyr	
				165					170)				175			
35	ATC	TTT	GCA	GAC	GCT	TAT	GCC	CAG	TAC	CTC	TGG	ATC	ACG	TTT	GAC	TTC	659
	Пе	Phe	Ala	Asp	Ala	Туг	Ala	Gln	Туг	Leu	Trp	Ile	Thr	Phe	Asp	Phe	
10			180					185					190				
••	TGC	AAC	ACT	CTT	CAA	GGC	TTT	TCC	ATC	CCA	TTT	. CCC	GCA	GCT	GAT	CTC	707
	Cys	Asn	Thr	Leu	Gln	Gly	Phe	Ser	Ile	Pro	Phe	Arg	Ala	Ala	Asp	Leu	
15		195					200					205			\sim		
	CTC (CTA (CAC /	AGT A	AAG (GCC 1	rcc /	AAC	CTT	CTC	TTG	GGC 1	rtt (GAC A	NGG 1	CC	755
	Leu l	Leu A	dis S	Ser I	Lys i	Ala S	Ser A	Asn	Leu	Leu	Leu	Gly I	Phe A	Asp A	arg S	Ser	
50	210					215					220					25	
	CAC (CCC A	AC A	AAG (CAG (CTG 1	rgg A	AAG	TCA	GAT	GAC	TTT (GC C	CAG A	CC T	`GG	803

	His	Pro	Asn	Lys	Gln	Leu	Trp	Lys	Ser	Asp	Asp	Phe	Gly	G1 n	Thr	Trp	
5					230					235					240		
	ATC	ATG	ТТА	CAG	GAA	CAT	GTC	AAG	TCC	TTT	TCT	TGG	GGA	ATT	GAT	CCC	851
	He	Met	He	Gln	Glu	His	Val	Lys	Ser	Phe	Ser	Trp	Gly	lle	Asp	Pro	
10				245					250					255			
					ААТ												899
	Туг	Asp	Lys	Pro	Asn	Thr	lle	Туг	Ile	Glu	Arg	His	Glu	Pro	Ser	Gly	
15			260					265					270				
	TAC	TCC	ACT	GTC	TTC	CGA	AGT	ACA	GAT	TTC	TTC	CAG	TCC	CGG	GAA	AAC	947
20	Tyr	Ser	Thr	Val	Phe	Arg	Ser	Thr	Asp	Phe	Phe	Gln	Ser	Arg	Glu	Asn	
		275					280					285					
	CAG	GAA	GTG	ATC	CTT	GAG	GAA	GTG	AGA	GAT	TTT	CAG	CTT	CGG	GAC	AAG	995
25	Gln	Glu	Val	He	Leu	Glu	Glu	Val	Arg	Asp	Phe	Gln	Leu	Arg	Asp	Lys	
	290					295					300					305	
	TAC	ATG	TTT	GCT	ACA	AAG	GTG	GTG	CAT	CTC	TTG	GGC	AGT	GAA	CAG	CAG	1043
30	Tyr	Met	Phe	Ala	Thr	Lys	Val	Val	His	Leu	Leu	Gly	Ser	Glu	Gln	Gln	
	•				310					315					320		
	TCT	тст	GTC	CAG	CTC	TGG	GTC	TCC	TTT	GGC	CGG	AAG	CCC	ATG	AGA	GCA	1091
35	Ser	Ser	Val	Gln	Leu	Trp	Val	Ser	Phe	Gly	Arg	Lys	Pro	Met	Arg	Ala	
				325					330					335			
40	GCC	CAG	TTT	GTC	ACA	AGA	CAT	CCT	ATT	AAT	GAA	TAT	TAC	ATC	GCA	GAT	1139
	Ala	Gln	Phe	Val	Thr	Arg	His	Pro	lle	Asn	Glu	Tyr	Tyr	He	Ala	Asp	
			340					345					350				
45	GCC	TCC	GAG	GAC	CAG	GTG	TTT	GTG	TGT	GTC	AGC	CAC	AGT	AAC	AAC	CGC	1187
	Ala	Ser	Glu	Asp	G1 n	Val	Phe	Val	Cys	Val	Ser	His	Ser	Asn	Asn	Arg	
		355					360					365					
50	ACC	AAT	TTA	TAC	ATC	TCA	GAG	GCA	GAG	GGG	СТG	AAG	TTC	TCC	CTG	TCC	1235
	Thr	Asn	Leu	Tyr	lle	Ser	Glu	Ala	Glu	Gly	Leu	Lys	Phe	Ser	Leu	Ser	

	370		375	380	•	385
	TTG GAG	AAC GTG CTC	TAT TAC AGC	CCA GGA GGG GC	C GGC AGT GAC	ACC 1283
5	Leu Glu	Asn Val Leu	Tyr Tyr Ser	Pro Gly Gly Ala	a Gly Ser Asp	Thr
		390		395	400	
10	TTG GTG	AGG TAT TTT	GCA AAT GAA	CCA TTT GCT GAG	C TTC CAC CGA	GTG 1331
	Leu Val	Arg Tyr Phe	Ala Asn Glu	Pro Phe Ala Ası	p Phe His Arg	Val
		405		410	415	
15	GAA GGA '	TTG CAA GGA	GTC TAC ATT	GCT ACT CTG AT	T AAT GGT TCT	ATG 1379
	Glu Gly	Leu Gln Gly	Val Tyr Ile	Ala Thr Leu Ile	e Asn Gly Ser	Met
20		420	425		430	
20	AAT GAG	GAG AAC ATG	AGA TCG GTC	ATC ACC TTT GA	C AAA GGG GGA	ACC 1427
		Glu Asn Met	Arg Ser Val	lle Thr Phe As	sp Lys Gly Gly	Thr
25	435		440	44		
				TTC ACG GGA TA		
	Trp Glu	Phe Leu Gln	Ala Pro Ala	Phe Thr Gly Ty	r Gly Glu Lys	lle
30	450		455	460		465
				TCC CTT CAT CT		
<i>35</i>	Asn Cys	Glu Leu Ser	Gln Gly Cys	Ser Leu His Le	eu Ala Gln Arg	Leu
		470		475	480	
				CGG AGA ATG CC		
40	Ser Gin		Leu Gin Leu	Arg Arg Met Pr		Lys
	040 000	485	0m2 1m2 1m2	490	495	
				GCC ACT GGC TC		
45				Ala Thr Gly Se		Asn
		500	505		510	100
50				ATC TCT AGC AG		
		SEL LAS IUL		Ile Ser Ser Se		Arg
	515		520	52	5	

	TGG	CGA	GAG	GCA	CTT	CCT	GGA	CCT	CAC	TAC	TAC	ACA	TGG	GGA	GAC	CAC	1715
5	Trp	Arg	Glu	Ala	Leu	Pro	Gly	Pro	His	Tyr	Tyr	Thr	Trp	Gly	Asp	His	
	530					535					540					545	
	GGC	GGA	ATC	ATC	ACG	GCC	ATT	GCC	CAG	GGC	ATG	GAA	ACC	AAC	GAG	CTA	1763
10	Gly	Gly	lle	Ile	Thr	Ala	Ile	Ala	Gln	Gly	Met	Glu	Thr	Asn	Glu	Leu	
					550					555	•				560		
	AAA	TAC	AGT	ACC	AAT	GAA	GGG	GAG	ACC	TGG	AAA	ACA	TTC	ATC	TTC	TCT	1811
15	Lys	Tyr	Ser	Thr	Asn	Glu	Gly	Glu	Thr	Trp	Lys	Thr	Phe	Ile	Phe	Ser	
				565					570					575			
20	GAG	AAG	CCA	GTG	TTT	GTG	TAT	GGC	CTC	CTC	ACA	GAA	CCT	GGG	GAG	AAG	1859
	Glu	Lys	Pro	Val	Phe	Val	Tyr	Gly	Leu	Leu	Thr	Glu	Pro	Gly	Glu	Lys	
			580					585					590				
25	AGC	ACT	GTC	TTC	ACC	ATC	TTT	GGC	TCG	AAC	AAA	GAG	AAT	GTC	CAC	AGC	1907
	Ser	Thr	Val	Phe	Thr	Ιle	Phe	Gly	Ser	Asn	Lys	Glu	Asn	Val	His	Ser	
		595					600					605					
30	TGG	CTG	ATC	СТС	CAG	GTC	AAT	GCC	ACG	GAT	GCC	TTG	GGA	GTT	CCC	TGC	1955
	Trp	Leu	lle	Leu	Gln	Val	Asn	Ala	Thr	Asp	Ala	Leu	Gly	Val	Pro	Cys	
35	610					615					620					625	
	ACA	GAG	AAT	GAC	TAC	AAG	CTG	TGG	TCA	CCA	TCT	GAT	GAG	CGG	GGG	AAT	2003
	Thr	Glu	Asn	Asp	Туг	Lys	Leu	Trp	Ser	Pro	Ser	Asp	Glu	Arg	Gly	Asn	
40					630					635					640		
	GAG	TGT	TTG	CTG	GGA	CAC	AAG	ACT	GTT	TTC	AAA	CGG	CGG	ACC	CCC	CAT	2051
	Glu	Cys	Leu	Leu	Gly	His	Lys	Thr	Val	Phe	Lys	Arg	Arg	Thr	Pro	His	
45				645					650					655			
	GCC	ACA	TGC	TTC	TAA	GGA	GAG	GAC	TTT	GAC	AGG	CCG	GTG	GTC	GTG	TCC	2099
	Ala	Thr	Cys	Phe	Asn	Gly	Glu	Asp	Phe	Asp	Arg	Pro	Val	Val	Val	Ser	
50			660					665					670				
	AAC	TGC	TCC	TGC	ACC	CGG	GAG	GAC	TAT	GAG	TGT	GAC	TTC	GGT	TTC	AAG	2147

	Asn	Cys	Ser	Cys	Thr	Arg	Glu	Asp	Туг	Glu	Cys	Asp	Phe	Gly	Phe	Lys	
5		675					680					685					
	ATG	AGT	GAA	GAT	TTG	TCA	TTA	GAG	GTT	TGT	GTT	CCA	GAT	CCG	GAA	TTT	2195
	Met	Ser	Glu	Asp	Leu	Ser	Leu	Glu	Val	Cys	Val	Pro	Asp	Pro	Glu	Phe	
10	690	,				695		•			700					705	•
	TCT	GGA	AAG	TCA	TAC	TCC	CCT	CCT	GTG	ССТ	TGC	CCT	GTG	GGT	тст	ACT	2243
	Ser	Gly	Lys	Ser	Туг	Ser	Pro	Pro	Val	Pro	Cys	Pro	Val	Gly	Ser	Thr	
15					710					715					720		
	TAC	AGG	AGA	ACG	AGA	GGC	TAC	CGG	AAG	ATT	TCT	GGG	GAC	ACT	TGT	AGC	2291
	Tyr	Arg	Arg	Thr	Arg	Gly	Tyr	Arg	Lys	Ιle	Ser	Gly	Asp	Thr	Cys	Ser	
20				725					730					735			
	GGA	GGA	GAT	GTT	GAA	GCG	CGA	СТG	GAA	GGA	GAG	CTG	GTC	ccc	TGT	CCC	2339
25	Gly	Gly	Asp	Val	Glu	Ala	Arg	Leu	Glu	Gly	Glu	Leu	Val	Pro	Cys	Pro	
			740					745					750				
	CTG	GCA	GAA	GAG	AAC	GAG	TTC	ATT	CTG	TAT	GCT	GTG	AGG	AAA	TCC	ATC	2387
30	Leu	Ala	Glu	Glu	Asn	Glu	Phe	lle	Leu	Tyr	Ala	Val	Arg	Lys	Ser	lle	
		755					760					765	i				
	TAC	CGC	ТАТ	GAC	CTG	GCC	TCG	GGA	GCC	ACC	GAG	CAG	TTG	CCT	CTC	ACC	2435
35	Tyr	Arg	Tyr	Asp	Leu	Ala	Ser	Gly	Ala	Thr	Glu	Gln	Leu	Pro	Leu	Thr	
	770					775	j				780)				785	
	GGG	СТА	CGG	GCA	GCA	GTG	GCC	CTC	GAC	777	GAC	TAT	GAG	CAC	CAAC	TGT	2483
40	Gly	Leu	Arg	Ala	Ala	. Val	Ala	Leu	Asp	Phe	. Asp	Туг	Glu	His	s Asr	Cys	
					790)				795	j .				800)	
45	TTO	ТАТ	`TGG	TCC	GAC	СТО	GCC	TTO	GAC	GTC	ATC	CAC	G CGC	CTO	TG1	TTG	2531
	Leu	туг	Trp	Ser	Asp	Leu	Ala	Leu	Asp	Val	I 1 e	e Glr	Arg	g Lei	ı Cys	s Leu	
				805	ō			-	810)				815	5 .		
50	AAT	r GGA	A AGC	ACA	GGC	G CAA	GAC	GTC	ATC	CATO	C AA1	r TC	r GGC	CTO	G GAC	G ACA	2579
	Asr	ı Gly	/ Ser	Thr	Gly	/ Glr	n Glu	ı Val	He	e Ile	e Ası	n Sei	Gly	/ Ĺei	ı Glu	ı Thr	

			820)				825					830)			
_	GTA	GAA	GCT	TTO	GCT	TTT	GAA	CCC	СТС	AGC	CAG	ста	СТТ	TAC	TGG	GTA	2627
5		Glu															
		835					840					845					
10	GAT	GCA	GGC	TTC	AAA	AAG	ATT	GAG	GTA	GCT	AAT	CCA	GAT	GGC	GAC	TTC	2675
	Asp	Ala	Gly	Phe	Lys	Lys	He	Glu	Val	Ala	Asn	Pro	Asp	Gly	Asp	Phe	
	850					855					860					865	
15	CGA	СТС	ACA	ATC	GTC	AAT	TCC	TCT	GTG	СТТ	GAT	CGT	CCC	AGG	GCT	CTG	2723
	Arg	Leu	Thr	Ile	Val	Asn	Ser	Ser	Val	Leu	Asp	Arg	Pro	Arg	Ala	Leu	
					870					875					880		
20	GTC	CTC	GTG	CCC	CAA	GAG	GGG	GTG	ATG	TTC	TGG	ACA	GAC	TGG	GGA	GAC	2771
	Val	Leu	Val	Pro	Gln	Glu	Gly	Val	Met	Phe	Trp	Thr	Asp	Trp	Gly	Asp	
25				885					890					895			
	CTG	AAG	ССТ	GGG	ATT	TAT	CGG	AGC	ААТ	ATG	GAT	GGT	тст	GCT	GCC	TAT	2819
	Leu	Lys	Pro	Gly	Ιlе	Туг	Arg	Ser	Asn	Met	Asp	Gly	Ser	Ala	Ala	Tyr	
30			900					905					910				
	CAC	CTG	GTG	TCT	GAG	GAT	GTG	AAG	TGG	CCC	ААТ	GGC	ATC	TCT	GTG	GAC	2867
	His	Leu	Val	Ser	Glu	Asp	Val	Lys	Тгр	Pro	Asn	Gly	lle	Ser	Val	Asp	
35		915					920		-			925					
	GAC	CAG	TGG	ATT	TAC	TGG	ACG	GAT	GCC	TAC	CTG	GAG	TGC	АТА	GAG	CGG	2915
10	Asp	Gln	Trp	Ile	Tyr	Trp	Thr	Asp	Ala	Tyr	Leu	Glu	Cys	Ιlе	Glu	Arg	
	930					935					940					945	
	ATC	ACG	TTC	AGT	GGC	CAG	CAG	CGC	тст	GTC	ATT	CTG	GAC	AAC	СТС	CCG	2963
1 5	Ile	Thr	Phe	Ser	Gly	G1 n	Gln	Arg	Ser	Val	Ile	Leu	Asp	Asn	Leu	Pro	
					950					95 5					960		
	CAC	CCC	TAT	GCC	ATT	GCT	GTC	TTT	AAG	AAT	GAA	ATC	TAC	TGG	GAT	GAC	3011
50	His	Pro	Tyr	Ala	Ile	Ala	Val	Phe	Lys	Asn	Glu	Ile	Туг	Trp	Asp	Asp	
				965					970					975			

	TGG	TCA	CAG	СТС	AGC	ATA	TTC	CGA	GCT	TCC	AAA	TAC	AGT	GGG	TCC	CAG	;	3059
5	Trp	Ser	Gln	Leu	Ser	Ile	Phe	Arg	Ala	Ser	Lys	Туг	Ser	Gly	Ser	Gln		
•			980					985					990					
	ATG	GAG	АТТ	CTG	GCA	AAC	CAG	CTC	ACG	GGG	СТС	ATG	GAC	ATG	AAG	ATT	- (3107
10	Met	Glu	lle	Leu	Ala	Asn	Gln	Leu	Thr	Gly	Leu	Met	Asp	Met	Lys	lle		
		995					1000					1005	5					
	TTC	TAC	AAG	GGG	AAG	AAC	ACT	GGA	AGC	AAT	GCC	TGT	GTG	ccc	AGG	CCA	. ;	3155
15	Phe	Туг	Lys	Gly	Lys	Asn	Thr	Gly	Ser	Asn	Ala	Cys	Val	Pro	Arg	Pro		
	1010)				1015	i				1020)				1025		
	TGC	AGC	CTG	CTG	TGC	CTG	CCC	AAG	GCC	AAC	AAC	AGT	AGA	AGC	TGC	AGG		3203
20	Cys	Ser	Leu	Leu	Cys	Leu	Pro	Lys	Ala	Asn	Asn	Ser	Arg	Ser	Cys	Arg		
					1030)				103	5				1040)		
25	TGT	CCA	GAG	GAT	GTG	TCC	AGC	AGT	GTG	CTT	CCA	TCA	GGG	GAC	CTG	ATG		3251
	Cys	Pro	Glu	Asp	Val	Ser	Ser	Ser	Val	Leu	Pro	Ser	Gly	Asp	Leu	Met		
	,			104	5				105	0				105	5			
30	TGT	GAC	TGC	CCT	CAG	GGC	TAT	CAG	СТС	AAG	AAC	AAT	ACC	TGT	GTC	AAA		3299
	Cys	Asp	Cys	Pro	Gln	Gly	Tyr	Gln	Leu	Lys	Asn	Asn	Thr	Cys	Val	Lys		
			106	0				106	5				107	0				
35	GAA	GAG	AAC	ACC	TGT	CTT	CGC	AAC	CAG	ТАТ	CCC	TGC	AGC	AAC	GGG	AAC		3347
	Glu	Glu	Asn	Thr	Cys	Leu	Arg	Asn	Gln	Tyr	Arg	Cys	Ser	Asn	Gly	Asn		
		107	5				108	0				108	35					
40	TGT	ATC	CAAC	AGC	ATT	TGG	TGG	TG1	GAC	TT	GAC	CAAC	GAC	TGT	` GGA	GAC		3395
	Cys	ille	e Asr	. Ser	lle	Trp	Trp	Cys	. Asp	Phe	e Ası	Ası	n Asp	Cys	Gly	Asp		
45	109	90				109	95				110	00				1105	Ś	
	ATC	G AGO	GA1	r GAC	G AGA	AAC	TGC	CC1	r acc	CAC	CATO	C TG	r GA	CTO	G GAC	ACC		3443
	Met	: Sei	Ası	o Glu	ı Arg	g Asr	ı Cys	Pro	Th:	r Th	r II	е Су	s As	Lei	ı Ası	Thr		
50					11	10				11	15				112	20		
	CAC	G TT	r cg	T TG	C CAC	G GAC	TC1	r GG	G AC	T TG	T AT	c cc	A CT	G TC	C TA	AAA 7		3491

	G	ln F	he	Arg	Cy:	s G1	n Gl	u Se	r Gl	y Ti	ir C	ys Il	e Pr	o Le	eu Se	er T	yr Lys	5
5					112	25				11	30				1	135		
	TC	ST G	AC	CTT	GA(G GA	T GA	C TG	T GG	A GA	C AA	C AG	T GA	T GA	A A(ST CA	AT TG1	3539
	Су	's A	sp	Leu	Glu	ı As	p As	р Су	s Gl	y As	p As	n Se	r As	p Gl	u Se	r Hi	s Cys	:
10				114						45					50			
																	C ATG	
	Gl	u M	et	His	Gln	Cys	s Ar	g Se	r As	p G1	u Ty	r As	n Cy	s Se	r Se	r Gl	y Met	
15			155					116					11					-
																	G GAC	3635
20	Cy:	s II	e i	Arg	Ser	Ser	Tri	o Val	Cys	s As _j	Gl	y Ası	Ası	ı Ası	р Су	s Ar	g Asp	
	11'						117					118					1189	5
	TG	G TC	Т (GAT	GAA	GCC	AAC	TG1	` ACC	GCC	CATO	C TA1	CAC	ACC	C TG	r ga	G GCC	3683
25	Trp	Se	r A	lsp	Glu	Ala	Asr	Cys	Thr	Ala	H	е Туг	His	Thr	. Cys	s Gli	u Ala	
						119	0				119	95				120	00	
	TCC	: AA	СТ	TC	CAG	TGC	CGA	AAC	GGG	CAC	TGC	ATC	ccc	CAG	CGC	TG(GCG	3731
30																	Ala	0.01
					1205					121					121			
-	TGT	GA	C G	GG (GAT	ACG	GAC	TGC	CAG	GAT	GGT	TCC	GAT	GAG			GTC	3779
35																	Val	0110
				220					1225					123				
10	AAC	TGT	r G	AG A	AG .	AAG	TGC	ААТ	GGA	TTC	CGC	TGC	CCA	AAC	GGC	ACT	TGC	3827
																	Cys	3021
		123						1240					1245		•		-,-	
5	ATC	CCA	TO	CC A	.GC /	AAA	САТ	TGT	GAT	GGT	CTG	CGT	GAT	TGC	TCT	GAT	GGC	3875
												Arg						00.0
	1250						1255					1260		•			1265	
0	тсс	GAT	GA	A C	AG (CAC '	TGC	GAG	CCC	СТС	TGT	ACG		TTC	ATG	GAC		3923
												Thr						0020
											-		_					

					1270					1275	j				1280		
5	GTG	TGT	AAG	AAC	CGC	CAG	CAG	TGC	CTG	TTC	CAC	TCC	ATG	GTC	TGT	GAC	3971
,	Val	Cys	Lys	Asn	Arg	Gln	Gln	Cys	Leu	Phe	His	Ser	Met	Val	Cys	Asp	
				1285	j				1290)				1295	5		
10	GGA	ATC	ATC	CAG	TGC	CGC	GAC	GGG	TCC	GAT	GAG	GAT	GCG	GCG	TTT	GCA	4019
	Gly	Ile	Ile	Gln	Cys	Arg	Asp	Gly	Ser	Asp	Glu	Asp	Ala	Ala	Phe	Ala	
			1300)				1305	5				1310)			
15	GGA	TGC	TCC	CAA	GAT	CCT	GAG	TTC	CAC	AAG	GTA	TGT	GAT	GAG	TTC	GGT	4067
	Gly	Cys	Ser	Gln	Asp	Pro	Glu	Phe	His	Lys	Val	Cys	Asp	Glu	Phe	Gly	
20		131	5				1320)				132	5				
20	TTC	CAG	TGT	CAG	AAT	GGA	GTG	TGC	ATC	AGT	TTG	ATT	TGG	AAG	TGC	GAC	4115
	Phe	Gln	Cys	Gln	Asn	Gly	Val	Cys	He	Ser	Leu	Ile	Trp	Lys	Cys	Asp	
25	133	0				133	5				134	0				1345	
	GGG	ATG	GAT	GAT	TGC	GGC	GAT	TAT	TCT	GAT	GAA	GCC	AAC	TGC	GAA	AAC	4163
	Gly	Met	Asp	Asp	Cys	Gly	Asp	Туг	Ser	Asp	Glu	Ala	. Asn	Cys	Glu	Asn	
30					135					135					136		
																` GAG	4211
	Pro	Thr	Glu	Ala	Pro	Asn	Cys	Ser	Arg	Tyr	Phe	Glr	n Phe	e Arg	g Cys	Glu	
35				136					137					137			
								,								C GAC	4259
40	Asr	Gly	, His	s Cys	: 11 <i>e</i>	Pro) Asr	ı Arg	Tr) Lys	s Cys	s Ası	p Ar	g Gl	u Asi	n Asp	
			138					138					13				
	TGT	r GG(G GAG	C TG(G TC	r Ga′	r GA(G AAC	G GA'	r TG	T GG	A GA	T TC	A CA	TAT	r ctt	4307
45	Cys	s Gl	y As	p Tr	Sei	r Ası	p Glu	ı Lys	s As	р Су:	s Gl	y As	p Se	r Hi	s II	e Leu	
		13					140						05				
	CC	CTT	СТС	G AC	r cc	T GG	G CC	C TC	C AC	G TG	т ст	G CC	C AA	T TA	C TA	C CGC	4355
50	Pr	o Ph	e Se	r Th	r Pr	o Gl	у Рг	o Se	r Th	г Су	s Le	u Pr	o As	n Ty	г Ту	r Arg	
	14	10				14	15				14	20				1425	

	TGC	AGC A	GT GO	G AC	C TG	C GT	G AT	G GA	C AC	с то	G G1	C T	GC G	AC G	GG TA	AC 4403
5	Cys	Ser S	er Gl	y Th	r Cy.	s Va	l Me	t As	p Th	г Тг	p Va	ıl C	ys As	sp G	ју Т	yr
				14					14						140	-
	CGA	GAT TO	GT GC	A GA	T GG	C TC1	r GA	C GA	G GA	A GC	C TG	c co	ст	rg C1	rr Go	CA 4451
10	Arg	Asp Cy	s Al	a As	p Gly	/ Ser	Ası	Gli	u Glt	u Al	а Су	s Pr	o Le	eu Le	u Al	a
			14					145						55		
	AAC (GTC AC	T GC	T GC	C TCC	CACT	, ccc	CAC	CAA	CT	r GG	G CG	A TG	T GA	.C CG	A 4499
15	Asn \	/al Th	r Al	a Ala	a Ser	Thr	Pro	Thr	Glr	ı Lei	ı Gl	y Ar	g Cy	s As	р Аг	g
			60				146						70			J
00	TTT (GAG TT	C GA	A TGC	CAC	CAA	CCG	AAC	ACC	TG1	` AT	r cc	C AA	С TG	G AA	G 4547
20	Phe G	ilu Ph	e Glu	ı Cys	His	Gln	Pro	Lys	Thr	Cys	: Ile	e Pr	o As	n Tr	p Ly	S
		475				148				,	148					
25	CGC T	GT GA	C GGC	CAC	CAA	GAT	TGC	CAG	GAT	GGC	CGC	GA	C GA	G GC	C AA	T 4595
	Arg C	ys Ası	Gly	His	Gln	Asp	Cys	Gln	Asp	Gly	Arg	Ası	o Gla	ı Ala	a Ası	n 1000
	1490				149					150					150	
30	TGC C	CC ACA	CAC	AGC	ACC	TTG	ACT	TGC	ATG	AGC	AGG	GAC	TTO	CAC		
	Cys Pi															
				1510					1515					152		'
35	GAG GA	C GGG	GAG	GCC	TGC	АТТ	GTG	CTC			CGC	TGC	GAC			4691
	Glu As															
40			u u		0,0	116	va!	Leu	Ser	Glu	Arg	Cvs	Asn	Glv	Phe	
40			1525		0,0	116		Leu 1530		Glu	Arg	Cys			Phe	
40			1525	5				1530)				153	5		
40	CTG GA	C TGC	1525 TCG	5 GAC	GAG	AGC	GAT	1530 GAA	AAG	GCC	TGC	AGT	153 GAT	5 GAG	TTG	4739
45		C TGC	1525 TCG Ser	5 GAC	GAG	AGC Ser	GAT Asp	1530 GAA Glu	AAG	GCC	TGC	AGT Ser	153 GAT Asp	5 GAG	TTG	4739
	CTG GA	C TGC p Cys 154	1529 TCG Ser	GAC Asp	GAG Glu	AGC Ser	GAT Asp 1545	1530 GAA Glu	AAG Lys	GCC Ala	TGC Cys	AGT Ser 1550	153 GAT Asp	5 GAG Glu	TTG Leu	4739
	CTG GA	C TGC p Cys 1540	1525 TCG Ser O	GAC Asp GTA	GAG Glu CAG	AGC Ser	GAT Asp 1545 CTT	1530 GAA Glu CAG	AAG Lys TGG	GCC Ala ACA	TGC Cys GCT	AGT Ser 1550 GAC	153 GAT Asp) TTC	5 GAG Glu TCT	TTG Leu GGG	4739
	CTG GALeu As	C TGC p Cys 1540	1525 TCG Ser O	GAC Asp GTA	GAG Glu CAG	AGC Ser AAT	GAT Asp 1545 CTT	1530 GAA Glu CAG	AAG Lys TGG	GCC Ala ACA	TGC Cys GCT Ala	AGT Ser 1550 GAC Asp	153 GAT Asp) TTC	5 GAG Glu TCT	TTG Leu GGG	4739
45	CTG GALeu As	C TGC p Cys 154 G TAC l Tyr	1525 TCG Ser O AAA Lys	GAC Asp GTA Val	GAG Glu CAG	AGC Ser AAT Asn	GAT Asp 1545 CTT	1530 GAA Glu CAG Gln	AAG Lys TGG Trp	GCC Ala ACA Thr	TGC Cys GCT Ala 1565	AGT Ser 1550 GAC Asp	153 GAT Asp) TTC Phe	5 GAG Glu TCT Ser	TTG Leu GGG Gly	4739

	Asp	Val	Thr	Leu	Thr	Trp	Met	Arg	Pro	Lys	Lys	Met	Pro	Ser	Ala	Ser	
_	1570)				1575	5				1580)				1585	
5	TGT	GTA	TAT	AAT	GTC	TAC	TAC	AGG	GTG	GTT	GGA	GAG	AGC	ATA	TGG	AAG	4883
	Cys	Val	Tyr	Asn	Val	Tyr	Tyr	Arg	Val	Val	Gly	Glu	Ser	Ile	Trp	Lys	
10					1590)				1595	5				1600)	
	ACT	CTG	GAG	ACC	CAC	AGC	AAT	AAG	ACA	AAC	ACT	GTA	TTA	AAA	GTC	TTG	4931
	Thr	Leu	Glu	Thr	His	Ser	Asn	Lys	Thr	Asn	Thr	Val	Leu	Lys	Val	Leu	
15				1605	ō				161	0				161	5		
	AAA	CCA	GAT	ACC	ACG	ТАТ	CAG	GTT	AAA	GTA	CAG	GTT	CAG	TGT	CTC	AGC	4979
	lys	Pro	Asp	Thr	Thr	Tyr	Gln	Val	Lys	Val	Gln	Val	Gln	Cys	Leu	Ser	
20			1620	0				1629	5				1630	0			
	AAG	GCA	CAC	AAC	ACC	AAT	GAC	TTT	GTG	ACC	CTG	AGG	ACC	CCA	GAG	GGA	5027
25	Lys	Ala	His	Asn	Thr	Asn	Asp	Phe	Val	Thr	Leu	Arg	Thr	Pro	Glu	Gly	
		1639	5				164	0				164	5				
	TTG	CCA	GAT	GCC	CCT	CGA	AAT	CTC	CAG	CTG	TCA	CTC	CCC	AGG	GAA	GCA	5075
30	Leu	Pro	Asp	Ala	Pro	Arg	Asn	Leu	Gln	Leu	Ser	Leu	Pro	Arg	Glu	Ala	
	165	0				165	5				166	0				1665	
	GAA	GGT	GTG	ATT	GTA	GGC	CAC	TGG	GCT	CCT	CCC	ATC	CAC	ACC	CAT	GGC	5123
35	Glu	Gly	Val	Ile	Val	Gly	His	Trp	Ala	Pro	Pro	Ile	His	Thr	His	Gly	
					167	0				167	5				168	30	
40	CTC	ATC	CGT	GAG	TAC	ATT	GTA	GAA	TAC	AGC	AGG	AGT	GGT	TCC	AAG	ATG	5171
40	Leu	Ile	Arg	Glu	Туг	He	Val	Glu	Туг	Ser	Arg	Ser	Gly	Ser	Lys	Met	
				168	5				169	0				169	5		
45	TGG	GCC	TCC	CAG	AGG	GCT	GCT	AGT	AAC	TTT	ACA	GAA	ATC	AAG	AAC	TTA	5219
	Trp	Ala	Ser	Gln	Arg	Ala	Ala	Ser	Asn	Phe	Thr	Glu	lle	Lys	Asn	Leu	
			170	0				170	5				171	0			
50	TTG	GTC	AAC	ACT	СТА	TAC	ACC	GTC	AGA	GTG	GCT	GCG	GTG	ACT	AGT	CGT	5267
	Leu	Val	Asn	Thr	Leu	Туг	Thr	Val	Arg	Val	Ala	Ala	Val	Thr	Ser	Arg	

		171	.5				172	0				172	5				
5	GGA	. ATA	GGA	AAC	TGG	AGC	GAT	TCT	AAA	TCC	A TT	ACC	ACC	АТА	AAA	GGA	5315
			Gly														
	173					173					174					1745	
10	AAA	GTG	ATC	CCA	CCA	CCA	GAT	ATC	CAC	ATT	GAC	AGC	тат	GGT	GAA	AAT	5363
	Lys	Val	Ile	Pro	Pro	Pro	Asp	Ile	His	lle	Asp	Ser	Tyr	Gly	Glu	Asn	
					175	0				175	5				176	0	
15	TAT	CTA	AGC	TTC	ACC	CTG	ACC	ATG	GAG	AGT	GAT	ATC	AAG	GTG	AAT	GGC	5411
	Туr	Leu	Ser	Phe	Thr	Leu	Thr	Met	Glu	Ser	Asp	Ile	Lys	Val	Asn	Gly	
				176	5				1770	0				177	5		
20	TAT	GTG	GTG	AAC	CTT	TTC	TGG	GCA	TTT	GAC	ACC	CAC	AAG	CAA	GAG	AGG	5459
	Tyr	Val	Val	Asn	Leu	Phe	Trp	Ala	Phe	Asp	Thr	His	Lys	Gln	Glu	Arg	
25			1780)				178	5				1790)			
	AGA	ACT	TTG	AAC	TTC	CGA	GGA	AGC	ATA	TTG	TCA	CAC	AAA	GTT	GGC	AAT	5507
	Arg	Thr	Leu	Asn	Phe	Arg	Gly	Ser	He	Leu	Ser	His	Lys	Val	Gly	Asn	
30		179	5				1800)				1805	5				
	CTG	ACA	GCT	CAT	ACA	TCC	ТАТ	GAG	ATT	TCT	GCC	TGG	GCC	AAG	ACT	GAC	5555
	Leu	Thr	Ala	His	Thr	Ser	Tyr	Glu	Ile	Ser	Ala	Trp	Ala	Lys	Thr	Asp	
35	1810)				1815	5				1820)				1825	
	TTG	GGG	GAT	AGC	CCT	CTG	GCA	TTT	GAG	CAT	GTT	ATG	ACC	AGA	GGG	GTT	5603
40	Leu	Gly	Asp	Ser	Pro	Leu	Ala	Phe	Glu	His	Val	Met	Thr	Arg	Gly	Val	
40					1830)				1835	5				1840)	
	CGC	CCA	CCT	GCA	CCT	AGC	CTC	AAG	GCC	AAA	GCC	ATC	AAC	CAG	ACT	GCA	5651
45	Arg	Pro	Pro	Ala	Pro	Ser	Leu	Lys	Ala	Lys	Ala	He	Asn	Gln	Thr	Ala .	
				1845	i				1850)				1855	·		
	GTG	GAA	TGT	ACC	TGG	ACC	GGC	CCC	CGG	ААТ	GTG	GTT	TAT	GGT	ТТА	TTC	5699
50	Val	Glu	Cys	Thr	Trp	Thr	Gly	Pro	Arg	Asn	Val	Val	Tyr	Gly	lle	Phe	
			1860)				1865	i				1870)			

	TAT	GCC	ACG	TCC	TTT	CTT	GAC	СТС	TAT	CGC	AAC	CCG	AAG	AGC	TTG	ACT ·	5747
_	Tyr	Ala	Thr	Ser	Phe	Leu	Asp	Leu	Tyr	Arg	Asn	Pro	Lys	Ser	Leu	Thr	
5		1875	5				1880)				1885	5				
	ACT	TCA	CTC	CAC	AAC	AAG	ACG	GTC	ATT	GTC	AGT	AAG	GAT	GAG	CAG	TAT	5795
10	Thr	Ser	Leu	His	Asn	Lys	Thr	Val	Ile	Val	Ser	Lys	Asp	Glu	Gln	Tyr	
	1890)				1899	5				190	0				1905	
	TTG	TTT	CTG	GTC	CGT	GTA	GTG	GTA	CCC	TAC	CAG	GGG	CCA	TCC	TCT	GAC	5843
15	Leu	Phe	Leu	Val	Arg	Val	Val	Val	Pro	Туг	Gln	Gly	Pro	Ser	Ser	Asp	
					1910)				1915	5				1920)	
	TAC	CTT	GTA	GTG	AAG	ATG	ATC	CCG	GAC	AGC	AGG	CTT	CCA	CCC	CGT	CAC	5891
20	Tyr	Val	Val	Val	Lys	Met	Пе	Pro	Asp	Ser	Arg	Leu	Pro	Pro	Arg	His	
				1929	<u>ว</u> ี				1930)				193	5		
25	СТС	CAT	GTG	GTT	CAT	ACG	GGC	AAA	ACC	TCC	GTG	GTC	ATC	AAG	TGG	GAA	5939
23	Leu	His	Val	Val	His	Thr	Gly	Lys	Thr	Ser	Val	Val	Ιlе	Lys	Trp	Glu	
			1940	0				194	5				195	0			
30	TCA	CCG	TAT	GAC	TCT	CCT	GAC	CAG	GAC	TTG	TTG	TAT	GCA	ATT	GCA	GTC	5987
	Ser	Pro	Tyr	Asp	Ser	Pro	Asp	G1 n	Asp	Leu	Leu	Tyr	Ala	He	Ala	Val	
		195	5				196	0				196	5				
35	AAA	GAT	CTC	ATA	AGA	AAG	ACT	GAC	AGG	AGC	TAC	: AAA	GTA	AAA	TCC	CGT	6035
	Lys	Asp	Leu	He	Arg	Lys	Thr	Asp	Arg	Ser	Туг	Lys	Val	Lys	Ser	Arg	
	197	0				197	5				198	80				1985	
40	AAC	AGC	ACT	GTG	GAA	TAC	ACC	CTT	AAC	AAG	110	GAC	CCT	GGC	GGG	AAA	6083
,	Asn	Ser	Thr	Val	Glu	Туг	Thr	Leu	Asn	Lys	Lei	Glu	Pro	Gly	Gly	Lys	
45					199	90				199	95				200	00	
	TAC	CAC	ATC	: ATT	GTC	CAA	CTO	GGC	AAC	ATC	G AGO	CAAA	GA1	TCC	AGC	ATA C	6131
	Туr	His	Ile	: Ile	. Val	Glr	Leu	Gly	Asn	Met	Ser	Lys	. Asp	Ser	Ser	lle	
50				200)5				201	.0				201	15		
	AAA	AT1	ACC	CACA	GT1	TC#	ATT A	TCA	GCA	CC1	r Gar	r gco	TTA	AAA	ATC	CATA	6179

•

	Lys	Ile	Thr	Thr	Val	Ser	Leu	Ser	Ala	Pro	Asp	Ala	Leu	Lys	ille	lle	
5			202	0				202	25				203	0			
	ACA	GAA	AAT	GAT	CAT	GT1	CTT	CTC	TTT	TGO	AAA	AGC	CTO	GCT	` TTA	AAG	6227
	Thr	Glu	Asn	Asp	His	Val	Leu	Leu	Phe	Trp	Lys	Ser	Leu	Ala	Leu	Lys	
10		203	5				204	0				204	5		-		
	GAA	AAG	CAT	TTT	AAT	` GAA	AGC	AGG	GGC	TAT	GAG	ATA	CAC	ATG	TTT	GAT	6275
	Glu	Lys	His	Phe	Asn	Gľu	Ser	Arg	Gly	Tyr	Glu	Ile	His	Met	Phe	Asp	
15	205	0				205	5				206	0				2065	
	AGT	GCC	ATG	AAT	ATC	ACA	GCT	TAC	CTT	GGG	AAT	ACT	ACT	GAC	ААТ	TTC	6323
20	Ser	Ala	Met	Asn	Ile	Thr	Ala	Tyr	Leu	Gly	Asn	Thr	Thr	Asp	Asn	Phe	
					2070)				207	5				208	כ	
	TTT	AAA	ATT	TCC	AAC	CTG	AAG	ATG	GGT	CAT	ААТ	TAC	ACG	TTC	ACC	GTC	6371
25	Phe	Lys	lle	Ser	Asn	Leu	Lys	Met	Gly	His	Asn	Tyr	Thr	Phe	Thr	Val	
				2085	j				2090)				209	5		
	CAA	GCA	AGA	TGC	CTT	TTT	GGC	AAC	CAG	ATC	TGT	GGG	GAG	CCT	GCC	ATC	6419
30	Gln	Ala	Arg	Cys	Leu	Phe	Gly	Asn	Gln	Ile	Cys	Gly	Glu	Pro	Ala	lle	
			2100)				2105	5				2110)			
35	CTG	CTG	TAC	GAT	GAG	CTG	GGG	TCT	GGT	GCA	GAT	GCA	TCT	GCA	ACG	CAG	6467
	Leu	Leu	Tyr	Asp	Glu	Leu	Gly	Ser	Glý	Ala	Asp	Ala	Ser	Ala	Thr	Gln	
		2115					2120)				2125	5				
40	GCT	GCC	AGA	TCT	ACG	GAT	GTT	GCT	GCT	GTG	GTG	GTG	CCC	ATC	TTA	TTC	6515
	Ala	Ala	Arg	Ser	Thr	Asp	Val	Ala	Ala	Val	Val	Val	Pro	lle	Leu	Phe	
	2130)				2135	j				2140)				2145	
45	CTG	ATA	CTG	CTG	AGC	CTG	GGG	GTG	GGG	TTT	GCC	ATC	CTG	TAC	ACG	AAG	6563
	Leu	lle	Leu	Leu	Ser	Leu	Gly	Val	Gly	Phe	Ala	lle	Leu	Туг	Thr	Lys	
					2150	;				2155	5				2160	l	
50	CAC	CGG	AGG	CTG	CAG	AGC	AGC	TTC	ACC	GCC	TTC	GCC	AAC	AGC	CAC	TAC	6611
	His	Arg	Arg	Leu	Gln	Ser	Ser	Phe	Thr	Ala	Phe	Ala	Asn	Ser	His	Tyr	

				2165	5				2170)				2179	5		
5	AGC	TCC	AGG	CTG	GGG	TCC	GCA	ATC	TTC	TCC	TCT	GGG	GAT	GAC	CTG	GGG	6659
	Ser	Ser	Arg	Leu	Gly	Ser	Ala	Ile	Phe	Ser	Ser	Gly	Asp	Asp	Leu	Gly	
			2180)				2185	5				2190	0			
10	GAA	GAT	GAT	GAA	GAT	GCC	CCT	ATG	ATA	ACT	GGA	TTT	TCA	GAT	GAC	GTC	6707
	Glu	Asp	Asp	Glu	Asp	Ala	Pro	Met	Ile	Thr	Gly	Phe	Ser	Asp	Asp	Val	
15		219	5				2200)				220	5				
	CCC	ATG	GTG	ATA	GCC	TGA	AAGA	GCT '	TTCC	rcac	ra G	AAAC	CAAA	T GG	TGTA	AATA	6762
	Pro	Met	Val	He	Ala												
20	2210)															
	TTT	TATT	rga 1	γαας	GATAC	GT TO	GATG	GTTT/	1 TT	TTAA	AAGA	TGC	ACTT	TGA (GTTG	CAATAT	6822
25	GTTA	ATTT	TTA 1	ratg(GGCC	AA A						•					6843
30																	
35																	
40																	
45																	
50																	
50																	
55																	

SEQUENCE LISTING

5	(1) GENERAL INFORMATION:	
10	 (i) APPLICANT: (A) NAME: KOWA CO., LTD. (B) STREET: 6-29, Nishiki 3-chome, Naka-ku, Nagoya-shi, (C) CITY: Aichi (E) COUNTRY: Japan (F) POSTAL CODE (ZIP): none 	
	(ii) TITLE OF INVENTION: NOVEL LDL RECEPTOR ANALOG PROTEIN AND TH	Œ
4-	(iii) NUMBER OF SEQUENCES: 7	
15 20	 (iv) COMPUTER READABLE FORM: (A) MEDIUM TYPE: Floppy disk (B) COMPUTER: IBM PC compatible (C) OPERATING SYSTEM: PC-DOS/MS-DOS (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO) 	
20	(2) INFORMATION FOR SEQ ID NO: 1:	
25	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 6639 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
	(ii) MOLECULE TYPE: cDNA to mRNA	
30		
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:	
	ATGGCGACAC GGAGCAGCAG GAGGGAGTCG CGACTCCCCT TCCTATTCAC CCTGGTCGCG	60
35	CTGCTGCCGC CCGGGGCTCT CTGCGAGGTG TGGACGCGGA CACTGCACGG CGGCCGCGCG	120
	CCCTTACCCC AGGAGCGGGG CTTCCGCGTG GTGCAGGGCG ACCCGCGCGA GCTGCGGCTG	180
	TGGGAGCGCG GGGATGCCAG GGGGGGAGC CGGGCGGACG AGAAGCCGCT CCGGAGGAGA	240
40	CGGAGCGCTG CCCTGCAGCC CGAGCCCATC AAGGTGTACG GACAGGTCAG CCTCAATGAT	300
	TCCCACAATC AGATGGTGGT GCACTGGCCC GGAGAGAAAA GCAACGTGAT CGTGGCCTTG	360
	GCCCGGGACA GCCTGGCGTT GGCCAGGCCC AGGAGCAGTG ATGTGTACGT GTCTTATGAC	420
1 5	TATGGAAAAT CATTCAATAA GATTTCAGAG AAATTGAACT TCGGCGCGGG AAATAACACA	480
	GAGGCTGTGG TGGCCCAGTT CTACCACAGC CCTGCGGACA ACAAACGGTA CATCTTCGCA	540
	GATGCCTACG CCCAGTATCT CTGGATCACG TTTGACTTCT GCAACACCAT CCATGGCTTT	600
50	TCCATCCCGT TCCGGGCAGC TGATCTCCTA CTCCACAGTA AGGCCTCCAA CCTTCTCCTG	660
	GGCTTCGACA GGTCTCACCC CAACAAGCAG CTGTGGAAGT CGGATGATTT TGGCCAGACC	720
	TGGATCATGA TTCAAGAACA CGTGAAGTCC TTTTCTTGGG GAATTGATCC CTATGACAAA	780

	CCAAACACCA	TCTACATCGA	ACGGCACGAA	CCTTCTGGCT	ACTCCACGGT	TTTCCGAAJT	840
	ACAGACTTCT	TCCAGTCCCG	GGAAAA.CCAG	GAAGTGATCT	TGGAGGAAGT	GAGAGACTTT	900
5	CAGCTTCGGG	ACAAGTACAT	GTTTGCTACA	AAGGTGGTGC	ATCTCTTGGG	CAGTCCACTG	960
	CAGTCTTCTG	TCCAGCTCTG	GGTCTCCTTT	GGCCGGAAGC	CCATGCGGGC	CGCCCAGTTT	1020
	GTTACAAGAC	ATCCTATCAA	CGAATATTAC	ATCGCGGATG	CCTCGGAGGA	CCAGGTGTTT	1080
10	GTGTGTGTCA	GTCACAGCAA	CAACCGCACC	AACCTCTACA	TCTCGGAGGC	AGAGGGCTTG	1140
	AAGTTCTCTC	TGTCCCTGGA	GAACGTGCTC	TACTACACCC	CGGGAGGGC	CGGCAGTGAC	1200
	ACCTTGGTGA	GGTACTTTGC	AAATGAACCG	TTTGCTGACT	TCCATCGTGT	GGAAGGGTTG	1260
15	CAGGGAGTCT	ACATTGCTAC	TCTGATTAAT	GGTTCTATGA	ATGAGGAGAA	CATGAGATCT	1320
	GTCATCACCT	TTGACAAAGG	GGGCACCTGG	GAATTTCTGC	AGGCTCCAGC	CTTCACGGGG	1380
	TATGGAGAGA	AAATCAACTG	TGAGCTGTCC	GAGGGCTGTT	CCCTCCACCT	GGCCCAGCGC	1440
20	CTCAGCCAGC	TGCTCAACCT	CCAGCTCCGG	AGGATGCCCA	TCCTGTCCAA	GGAGTCGGCG	1500
	CCTGGCCTCA	TCATTGCCAC	GGGCTCAGTG	GGAAAGAACT	TGGCTAGCAA	GACAAACGTG	1560
	TACATCTCTA	GCAGTGCTGG	AGCCAGGTGG	CGAGAGGCAC	TTCCTGGACC	TCACTACTAT	1620
25	ACATGGGGAG	ACCATGGCGG	CATCATCATG	GCCATTGCCC	AAGGCATGGA	AACCAACGAA	1680
	CTGAAGTACA	GTACCAACGA	AGGGGAGACC	TGGAAAGCCT	TCACCTTCTC	TGAGAAGCCC	1740
	GTGTTTGTGT	ATGGGCTCCT	CACGGAACCC	GGCGAGAAGA	GCACGGTCTT	CACCATCTTT	1800
30	GGCTCCAACA	AGGAGAACGT	GCACAGCTGG	CTCATCCTCC	AGGTCAATGC	CACAGACGCC	1860
	CTGGGGGTTC	CTTGCACAGA	GAACGACTAC	AAGCTCTGGT	CACCATCTGA	TGAGCGGGGG	1920
	AATGAGTGTT	TGCTTGGACA	CAAGACTGTT	TTCAAACGGA	GGACCCCGCA	CGCCACATGC	1980
35	TTTAACGGAG	AAGACTTTGA	CAGGCCGGTG	GTTGTGTCCA	ACTGCTCCTG	CACCCGGGAG	2040
	GACTATGAGT	GTGACTTTGG	CTTCCGGATG	AGTGAAGACT	TGGCATTAGA	GGTGTGTGTT	2100
	CCAGATCCAG	GATTTTCTGG	AAAGTCCTCC	CCTCCAGTGC	CTTGTCCCGT	GGGCTCTACG	2160
40	TACAGGCGAT	CAAGAGGCTA	CCGGAAGATT	TCTGGGGACA	CCTGTAGTGG	AGGAGATGTT	2220
40	GAGGCACGGC	TAGAAGGAGA	GCTGGTCCCC	TGTCCCCTGG	CAGAAGAGAA	CGAGTTCATC	2280
	CTGTACGCCA	CGCGCAAGTC	CATCCACCGC	TATGACCTGG	CTTCCGGAAC	CACGGAGCAG	2340
	TTGCCCCTCA	CTGGGTTGCG	GGCAGCAGTG	GCCCTGGACT	TTGACTATGA	GCACAACTGC	2400
45	CTGTATTGGT	CTGACCTGGC	CTTGGACGTC	ATCCAGCGCC	TCTGTTTGAA	CGGGAGTACA	2460
	GGACAAGAGG	TGATCATCAA	CTCTGACCTG	GAGACGGTAG	AAGCTTTGGC	TTTTGAACCC	2520
	CTCAGCCAAT	TACTTTACTG	GGTGGACGCA	GGCTTTAAAA	AGATCGAGGT	AGCCAATCCA	2580
50	GATGGTGACT	TCCGACTCAC	CGTCGTCAAT	TCCTCGGTGC	TGGATCGGCC	CCGGGCCCTG	2640
	CTCCTTCTCC	CCCNACNACC	CATCATCTTC	ፐርርል ርርርልርፕ	GGGGAGACCT	GAAGCCTGGG	2700

	ATTTATCGGA	GCAACATGGA	CGGATCTGCC	GCCTATCGCC	TCGTG ICGGA	GG.ATGTGAAG	2760
5	TGGCCCAATG	GCATTTCCGT	GGACGATCAG	TGGATCTACT	GGACGGATGC	CTACCTGGAC	2820
J	TGCATTGAGC	GCATCACGTT	CAGCGGCCAG	CAGCGCTCCG	TCATCCTGGA	CAGACTCCCG	2880
	CACCCCTATG	CCATTGCTGT	CTTTAAGAAT	GAGATTTACT	GGGATGACTG	GTCACAGCTC	2940
	AGCATATTCC	GAGCTTCTAA	GTACAGCGGG	TCCCAGATGG	AGATTCTGGC	CAGCCAGCTC	3000
10	ACGGGGCTGA	TGGACATGAA	GATCTTCTAC	AAGGGGAAGA	ACACAGGAAG	CAATGCGTGT	3060
	GTACCCAGGC	CGTGCAGCCT	GCTGTGCCTG	CCCAGAGCCA	ACAACAGCAA	AAGCTGCAGG	3120
	TGTCCAGATG	GCGTGGCCAG	CAGTGTCCTC	CCTTCCGGGG	ACCTGATGTG	TGACTGCCCT	3180
15	AAGGGCTACG	AGCTGAAGAA	CAACACGTGT	GTCAAAGAAG	AAGACACCTG	TCTGCGCAAC	3240
	CAGTACCGCT	GCAGCAACGG	GAACTGCATC	AACAGCATCT	GGTGGTGCGA	TTTCGACAAC	3300
	GACTGCGGAG	ACATGAGCGA	CGAGAAGAAC	TGCCCTACCA	CCATCTGCGA	CCTGGACACC	3360
20	CAGTTCCGTT	GCCAGGAGTC	TGGGACGTGC	ATCCCGCTCT	CCTACAAATG	TGACCTCGAG	3420
	GATGACTGTG	GGGACAACAG	TGACGAAAGG	CACTGTGAAA	TGCACCAGTG	CCGGAGCGAC	3480
	GAATACAACT	GCAGCTCGGG	CATGTGCATC	CGCTCCTCCT	GGGTGTGCGA	CGGGGACAAC	3540
25	GACTGCAGGG	ACTGGTCCGA	CGAGGCCAAC	TGCACAGCCA	TCTATCACAC	CTGTGAGGCC	3600
	TCCAACTTCC	AGTGCCGCAA	CGGGCACTGC	ATCCCCCAGC	GGTGGGCGTG	TGACGGCGAC	3660
	GCCGACTGCC	AGGATGGCTC	TGATGAGGAT	CCAGCCAACT	GTGAGAAGAA	GTGCAACGGC	3720
30	TTCCGCTGCC	CGAACGGCAC	CTGCATTCCC	TCCACCAAGC	ACTGTGACGG	CCTGCACGAT	3780
	TGCTCGGACG	GCTCCGACGA	GCAGCACTGC	GAGCCCCTGT	GTACACGGTT	CATGGACTTC	3840
	GTGTGTAAGA	ACCGCCAGCA	GTGCCTCTTC	CACTCCATGG	TGTGCGATGG	GATCATCCAG	3900
35	TGCCGTGACG	GCTCCGACGA	GGACCCAGCC	TTTGCAGGAT	GCTCCCGAGA	CCCCGAGTTC	3960
	CACAAGGTGT	GCGATGAGTT	CGGCTTCCAG	TGTCAGAACG	GCGTGTGCAT	CAGCTTGATC	4020
	TGGAAGTGCG	ACGGGATGGA	TGACTGCGGG	GACTACTCCG	ACGAGGCCAA	CTGTGAAAAC	4080
40	CCCACAGAAG	CCCCCAACTG	CTCCCGCTAC	TTCCAGTTCC	GGTGTGACAA	TGGCCACTGC	4140
	ATCCCCAACA	GGTGGAAGTG	TGACAGGGAG	AATGACTGTG	GGGACTGGTC	CGACGAGAAG	4200
	GACTGTGGAG	ATTCACATGT	ACTTCCGTCT	ACGACTCCTG	CACCCTCCAC	GTGTCTGCCC	4260
4 5	AATTACTACC	GCTGCGGCGG	GGGGGCCTGC	GTGATAGACA	CGTGGGTTTG	TGACGGGTAC	4320
45	CGAGATTGCG	CAGATGGATC	CGACGAGGAA	GCCTGCCCCT	CGCTCCCCAA	TGTCACTGCC	4380
	ACCTCCTCCC	CCTCCCAGCC	TGGACGATGC	GACCGATTTG	AGTTTGAGTG	CCACCAGCCA	4440
	AAGAAGTGCA	TCCCTAACTG	GAGACGCTGT	GACGGCCATC	AGGATTGCCA	GGATGGCCAG	4500
50	GACGAGGCCA	ACTGCCCCAC	TCACAGCACC	TTGACCTGCA	TGAGCTGGGA	GTTCAAGTGT	4560
	GAGGATGGCG	AGGCCTGCAT	CGTGCTGTCA	GAACGCTGCG	ACGGCTTCCT	GGACTGCTCA	4620

	GATGAGAGCG	ACGAGAAGGC	CTGCAGTGAT	GAGTTAACTG	TATACAAAGT	ACAGAATCTT	4680
	CAGTGGACAG	CTGACTTCTC	TGGGAATGTC	ACTTTGACCT	GGATGCGGCC	CAAAAAAATG	4740
5	CCCTCTGCTG	CTTGTGTATA	CAACGTGTAC	TATAG.AGTTG	TTGGAGAGAG	CATATGGAAG	4800
	ACTCTGGAGA	CTCACAGCAA	TAAGACAAAC	ACTGTATTAA	AAGTGTTGAA	ACCAGATACC	4860
	ACCTACCAGG	TTAAAGTGCA	GGTTCAGTGC	CTGAGCAAGG	TGCACAACAC	CAATGACTTT	4920
10	GTGACCTTGA	GAACTCCAGA	GGGATTGCCA	GACGCCCCTC	AGAACCTCCA	GCTGTCGCTC	4980
	CACGGGGAAG	AGGAAGGTGT	GATTGTGGGC	CACTGGAGCC	CTCCCACCCA	CACCCACGGC	5040
	CTCATTCGCG	AATACATTGT	AGAGTATAGC	AGGAGTGGTT	CCAAGGTGTG	GACTTCAGAA	5100
15	AGGGCTGCTA	GTAACTTTAC	AGAAATAAAG	AACTTGTTGG	TCAACACCCT	GTACACCGTC	5160
	AGAGTGGCTG	CGGTGACGAG	TCGTGGGATA	GGAAACTGGA	GCGATTCCAA	ATCCATTACC	5220
	ACCGTGAAAG	GAAAAGCGAT	CCCGCCACCA	AATATCCACA	TTGACAACTA	CGATGAAAAT	5280
20	TCCCTGAGTT	TTACCCTGAC	CGTGGATGGG	AACATCAAGG	TGAATGGCTA	TGTGGTGAAC	5340
	CTTTTCTGGG	CATTTGACAC	CCACAAACAA	GAGAAGAAAA	CCATGAACTT	CCAAGGGAGC	5400
	TCAGTGTCCC	ACAAAGTTGG	CAATCTGACA	GCACAGACGG	CCTATGAGAT	TTCCGCCTGG	5460
25	GCCAAGACTG	ACTTGGGCGA	TAGTCCTCTG	TCATTTGAGC	ATGTCACGAÇ	CAGAGGGGTT	5520
	CGCCCACCTG	CTCCTAGCCT	CAAGGCCAGG	GCTATCAATC	AGACTGCAGT	GGAATGCACC	5580
	TGGACAGGCC	CCAGGAATGT	GGTGTATGGC	ATTTTCTATG	CCACATCCTT	CCTGGACCTC	5640
30	TACCGCAACC	CAAGCAGCCT	GACCACGCCG	CTGCACAACG	CAACCGTGCT	CGTCGGTAAG	5700
50	GATGAGCAGT	ATCTGTTTCT	GGTCCGGGTG	GTGATGCCCT	ACCAAGGGCC	GTCCTCGGAC	5760
	TACGTGGTCG	TGAAGATGAT	CCCGGACAGC	AGGCTTCCTC	CCCGGCACCT	GCATGCCGTT	5820
	CACACCGGCA	AGACCTCGGC	CGTCATCAAG	TGGGAGTCGC	CCTACGACTC	TCCTGACCAG	5880
35	GACCTGTTCT	ATGCGATCGC	AGTTAAAGAT	CTGATACGAA	AGACGGACCG	GAGCTACAAA	5940
	GTCAAGTCCC	GCAACAGCAC	CGTGGAGTAC	ACCCTGAGCA	AGCTGGAGCC	CGGAGGGAAA	6000
	TACCACGTCA	TTGTGCAGCT	GGGGAACATG	AGCAAAGATG	CCAGTGTGAA	GATCACCACC	6060
40	GTTTCGTTAT	CGGCACCCGA	TGCCTTAAAA	ATCATAACAG	AAAATGACCA	CGTCCTTCTC	6120
	TTCTGGAAAA	GTCTAGCTCT	AAAGGAAAAG	TATTTTAACG	AAAGCAGGGG	CTACGAGATA	6180
	CACATGTTTG	ATAGCGCCAT	GAATATCACC	GCATACCTTG	GGAATACTAC	TGACAATTTC	624
45	TTTAAAATTT	CCAACCTGAA	GATGGGTCAC	AATTACACAT	TCACGGTCCA	GGCACGATGC	6300
	CTTTTGGGCA	GCCAGATCTG	CGGGGAGCCT	GCCGTGCTAC	TGTATGATGA	GCTGGGGTCT	636
	GGTGGCGATG	CGTCGGCGAT	GCAGGCTGCC	AGGTCTACTG	ATGTCGCCGC	CGTGGTGGTG	642
50	CCCATCCTGT	TTCTGATACT	GCTGAGCCTG	GGGGTCGGGT	TTGCCATCCT	GTACACGAAG	648
	ON THOROUGH CO.	TO 03 03 003 0	OMMON COCCO		CCCS CTS CSC	CTCCAGACTC	654

	GGCTCCGC	CA T	CTTC	TCCT	C TG	GGGA	TGAC	TTG	GGGG.	AGG .	ATGA	TGAA	GA Ţ	GCTC	CTAT	G	۴600
	ATCACTGG	AT T	TTCG	GACG.	A CG	TCCC	CATG	GTG.	ATAG	CC							6639
5	(2) INFO	RMAT	ION :	FOR	SEQ	ID N	0: 2	:									
10	(i)	(B (C	UENC:) LEI) TY:) STI) TO:	NGTH PE: RAND	: 22 amin EDNE	13 a: o ac: SS:	mino id	S: aci	ds								-
	(ii)	MOL	ECULI	E TY	PE:	prot	ein										
15	(xi)	SEQ	UENCI	E DE	SCRI	PTIO	N: S	EQ I	D NO	: 2:							
	Met 1	Ala	Thr	Arg	Ser 5	Ser	Arg	Arg	Glu	Ser 10	Arg	Leu	Pro	Phe	Leu 15	Phe	
20	Thr	Leu	Val	Ala 20	Leu	Leu	Pro	Pro	Gly 25	Ala	Leu	Cys	Glu	Val 30	Trp	Thr	
	Arg	Thr	Leu 35	His	Gly	Gly	Arg	Ala 40	Pro	Leu	Pro	Gln	Glu 45	Arg	Gly	Phe	
25	Arg	Val 50	Val	Gln	Gly	Asp	Pro 55	Arg	Glu	Leu	Arg	Leu 60	Trp	Glu	Arg	Gly	
	Asp 65	Ala	Arg	Gly	Ala	Ser 70	Arg	Ala	Asp	Glu	Lys 75	Pro	Leu	Arg	Arg	Arg 80	
30	Arg	Ser	Ala	Ala	Leu 85	Gln	Pro	Glu	Pro	Ile 90	Lys	Val	Tyr	Gly	Gln 95	Val	
	Ser	Leu	Asn	Asp 100	Ser	His	Asn	Gln	Met 105	Val	Val	His	Trp	Ala 110	Gly	Glu	
35		Ser	115					120					125				
		Pro 130					135					140					
40	145	Asn				150					155					160	
		Ala			165					170					175		
45		Ile		180					185					190			
		Сув	195					200					205			_	·
50		Leu 210					215					220					
	Ser 225	His	Pro	Asn	Lys	Gln 230	Leu	Trp	Lys	Ser	Asp 235	Asp	Phe	Gly	Gln	Thr 240	
55																	

	Trp	Ile	Met	Ile	Gln 245	Glu	His	Val	üys	Ser 250	₽he	Ser	Trp	Gly	Ile 255	ga.4
5	Pro	Tyr	Asp	Lys 260	Pro	Asn	Thr	Ile	Tyr 265	Ile	Glu	Arg	His	Glu 270	Pro	Ser
	Gly	Tyr	Ser 275	Thr	Val	Phe	Arg	Ser 280	Thr	Asp	Phe	Phe	Gln 285	Ser	Arg	Glu
10	Asn	Gln 290	Glu	Val	Ile	Leu	Glu 295	Glu	Val	Arg	Asp	Phe 300	Gln	Leu	Arg	Asp
	Lys 305	Tyr	Met	Phe	Ala	Thr 310	Lys	Val	Val	His	Leu 315	Leu	Gly	Ser	Pro	Leu 320
15	Gln	Ser	Ser	Val	Gln 325	Leu	Trp	Val	Ser	Phe 330	Gly	Arg	Lys	Pro	Met 335	Arg
	Ala	Ala	Gln	Phe 340	Val	Thr	Arg	His	Pro 345	Ile	Asn	Glu	Tyr	Tyr 350	Ile	Ala
20	Asp	Ala	Ser 355	Glu	Asp	Gln	Val	Phe 360	Val	Cys	Val	Ser	His 365	Ser	Asn	Asn
	Arg	Thr 370	Asn	Leu	Tyr	Ile	Ser 375	Glu	Ala	Glu	Gly	Leu 380	Lys	Phe	Ser	Leu
25	Ser 385	Leu	Glu	Asn	Val	Leu 390	Tyr	Tyr	Thr	Pro	Gly 395	Gly	Ala	Gly	Ser	Asp
	Thr	Leu	Val	Arg	Tyr 405	Phe	Ala	Asn	Glu	Pro 410	Phe	Ala	Asp	Phe	His 415	Arg
30	Val	Glu	Gly	Leu 420	Gln	Gly	Val	Tyr	Ile 425	Ala	Thr	Leu	Ile	Asn 430	Gly	Ser
	Met	Asn	Glu 435	Glu	Asn	Met	Arg	Ser 440	Val	Ile	Thr	Phe	Asp 445	Lys	Gly	Gly
35	Thr	Trp 450	Glu	Phe	Leu	Gln	Ala 455	Pro	Ala	Phe	Thr	Gly 460	Tyr	Gly	Glu	Lys
	Ile 465	Asn	Суз	Glu	Leu	Ser 470	Glu	Gly	Cys	Ser	Leu 475	His	Leu	Ala	Gln	Arg 480
40	Leu	Ser	Gln	Leu	Leu 485	Asn	Leu	Gln	Leu	Arg 490	Arg	Met	Pro	Ile	Leu 495	Ser
	Lys	Glu	Ser	Ala 500	Pro	Gly	Leu	Ile	11e 505	Ala	Thr	Gly	Ser	Val 510	Gly	Lys
45	Asn	Leu	Ala 515	Ser	Lys	Thr	Asn	Val 520	Tyr	Ile	Ser	Ser	Ser 525	Ala	Gly	Ala
45	Arg	Trp 530	Arg	Glu	Ala	Leu	Pro 535	Gly	Pro	His	Tyr	Tyr 540	Thr	Trp	Gly	Asp
	His 545	Gly	Gly	Ile	Ile	Met 550	Ala	Ile	Ala	Gln	Gly 555	Met	Glu	Thr	Asn	Glu 560
50	Leu	Lys	Tyr	Ser	Thr 565	Asn	Glu	Gly	Glu	Thr 570	Trp	Lys	Ala	Phe	Thr 575	Phe

	Ser	Glu	Lys	Pro 580	Val	Phe	Val	Tyr	Gly 585	Leu	Leu	Thr	Glu	Pro 590	Gly	Gīu
5	Lys	Ser	Thr 595	Val	Phe	Thr	Ile	Phe 600	Gly	Ser	Asn	Lys	Glu 605	Asn	Val	His
	Ser	Trp 610	Leu	Ile	Leu	Gln	Val 615	Asn	Ala	Thr	Asp	Ala 620	Leu	Gly	Val	Pro
10	Cys 625	Thr	Glu	Asn	Asp	Tyr 630	Lys	Leu	Trp	Ser	Pro 635	Ser	Asp	Glu	Arg	Gly 640
	Asn	Glu	Сув	Leu	Leu 645	Gly	His	Lys	Thr	Val 650	Phe	Lys	Arg	Arg	Thr 655	Pro
15	His	Ala	Thr	Cys 660	Phe	Asn	Gly	Glu	Asp 665	Phe	Asp	Arg	Pro	Val 670	Val	Val
	Ser	Asn	Cys 675	Ser	Суз	Thr	Arg	Glu 680	Asp	Tyr	Glu	Cys	Asp 685	Phe	Gly	Phe
20	Arg	Met 690	Ser	Glu	Asp	Leu	Ala 695	Leu	Glu	Val	Cys	Val 700	Pro	Asp	Pro	Gly
	Phe 705	Ser	Gly	Lys	Ser	Ser 710	Pro	Pro	Val	Pro	Cys 715	Pro	Val	Gly	Ser	Thr 720
25	Tyr	Arg	Arg	Ser	Arg 725	Gly	Tyr	Arg	Lys	Ile 730	Ser	Gly	Asp	Thr	Cys 735	Ser
	Gly	Gly	Asp	Val 740	Glu	Ala	Arg	Leu	Glu 745	Gly	Glu	Leu	Val	Pro 750	Cys	Pro
30	Leu	Ala	Glu 755	Glu	Asn	Glu	Phe	Ile 760	Leu	Tyr	Ala	Thr	Arg 765	Lys	Ser	Ile
	His	Arg 770	Tyr	Asp	Leu	Ala	Ser 775	Gly	Thr	Thr	Glu	Gln 780	Leu	Pro	Leu	Thr
35	Gly 785	Leu	Arg	Ala	Ala	Val 790	Ala	Leu	Asp	Phe	Asp 795	Tyr	Glu	His	Asn	Суs 800
	Leu	Tyr	Trp	Ser	Asp 805	Leu	Ala	Leu	Asp	Val 810	Ile	Gln	Arg	Leu	Cys 815	Leu
40	Asn	Gly	Ser	Thr 820	Gly	Gln	Glu	Val	Ile 825	Ile	Asn	Ser	Asp	Leu 830	Glu	Thr
	Val	Glu	Ala 835	Leu	Ala	Phe	Glu	Pro 840	Leu	Ser	Gln	Leu	Leu 845	Tyr	Trp	Val
4E	Asp	Ala 850	Gly	Phe	Lys	Lys	Ile 855	Glu	Val	Ala	Asn	Pro 860	Asp	Gly	Asp	Phe
4 5	Arg 865	Leu	Thr	Val	Val	Asn 870	Ser	Ser	Val	Leu	Asp 875	Arg	Pro	Arg	Ala	Leu 880
	Val	Leu	Val	Pro	Gln 885	Glu	Gly	Ile	Met	Phe 890	Trp	Thr	Asp	Trp	Gly 895	Asp
50	Leu	Lys	Pro	Gly 900	Ile	Tyr	Arg	Ser	Asn 905	Met	Asp	Gly	Ser	Ala 910	Ala	Tyr
	Arg	Leu	Val	Ser	Glu	Asp	Val	Lys	Trp	Pro	Asn	Gly	Ile	Ser	Val	Asp

			915					920					925			
5	Asp	Gln 930	Trp	Ile	Tyr	Trp	Thr 935	Asp	Ala	Tyr	Leu	Asp 940	Cys	Ile	Glu	Arg
	Ile 945	Thr	Phe	Ser	Gly	Gln 950	Gln	Arg	Ser	Val	Ile 955	Leu	Asp	Arg	Leu	Pro 960
10	His	Pro	Tyr	Ala	Ile 965	Ala	Val	Phe	Lys	Asn 970	Glu	Ile	Tyr	Trp	Asp 975	Asp
10	Trp	Ser	Gln	Leu 980	Ser	Ile	Phe	Arg	Ala 985	Ser	Lys	Tyr	Ser	Gly 990	Ser	Gln
	Met	Glu	Ile 995	Leu	Ala	Ser	Gln	Leu 100	Thr	Gly	Leu	Met	Asp		Lys	Ile
15	Phe	Tyr 101	Lys)	Gly	Lys	Asn	Thr 1019	Gly	Ser	Asn	Ala	Cys 102	Val	Pro	Arg	Pro
	Cys 1025	Ser 5	Leu	Leu	Cys	Leu 1030	Pro	Arg	Ala	Asn	Asn 1035		Lys	Ser	Cys	Arg
20	Cys	Pro	Asp	Gly	Val 1049	Ala	Ser	Ser	Val	Leu 1050	Pro	Ser	Gly	Asp	Leu 105	Met 5
	Cys	Asp	Cys	Pro 1060	Lys)	Gly	Tyr	Glu	Leu 1065		Asn	Asn	Thr	Cys 107		Lys
25	Glu	Glu	Asp 1075	Thr	Cys	Leu	Arg	Asn 1080	Gln	Tyr	Arg	Cys	Ser 1089		Gly	Asn
	Cys	Ile 1090	Asn	Ser	Ile	Trp	Trp 1095	Сув	Asp	Phe	Asp	Asn 1100		Суѕ	Gly	Asp
30	Met 1105	Ser	Asp	Glu	Lys	Asn 1110	Cys)	Pro	Thr	Thr	Ile 1115		Asp	Leu	Asp	Thr 1120
	Gln	Phe	Arg	Cys	Gln 1125	Glu	Ser	Gly	Thr	Сув 1130		Pro	Leu	Ser	Tyr 1135	
35	Cys	Asp	Leu	Glu 1140	Asp	Asp	Cys	Gly	Asp 1145		Ser	Asp	Glu	Arg 1150		Cys
	Glu	Met	His 1155	Gln	Сув	Arg	Ser	Asp 1160	Glu	Tyr	Asn	Cys	Ser 1165		Gly	Met
40	Cys	Ile 1170	Arg	Ser	Ser	Trp	Val 1175	Cys	Asp	Gly	Asp	Asn 1180		Cys	Arg	Asp
	Trp 1185	Ser	Asp	Glu	Ala	Asn 1190	Cys	Thr	Ala	Ile	Tyr 1195		Thr	Cys	Glu	Ala 1200
45	Ser	Asn	Phe	Gln	Cys 1205	Arg	Asn	Gly	His	Cys 1210	Ile	Pro	Gln	Arg	Trp 1215	
	Cys	Asp	Gly	Asp 1220	Ala	Asp	Cys	Gln	Asp 1225	Gly	Ser	Asp	Glu	Asp 1230		Ala
50	Asn	Cys	Glu 1235	Lys	Lys	Cys	Asn	Gly 1240		Arg	Сув	Pro	Asn 1245		Thr	Cys
	Ile	Pro 1250	Ser	Thr	Lys	His	Cys 1255		Gly	Leu	His	Asp 1260		Ser	qaA	Gly

	Ser 1265	Asp	Glu	Gln	His	Cys 1270	Glu	Pro	Leu	Cys	Thr 1275		Phe	Met	Asp	Pile 1280
5	Val	Сув	Lys	Asn	Arg 1285		Gln	Cys	Leu	Phe 1290		Ser	Met	Val	Cys 1295	-
	Gly	Ile	Ile	Gln 1300	Cys	Arg	Asp	Gly	Ser 1305		Glu	Asp	Pro	Ala 1310		Ala
10	Gly	Cys	Ser 1315		Asp	Pro	Glu	Phe 1320		Lys	Val	Cys	Asp 1325		Phe	Gly
	Phe	Gln 1330		Gln	Asn	Gly	Val 1335		Ile	Ser	Leu	Ile 1340		Lys	Cys	Asp
15	Gly 1345		Asp	Asp	Сув	Gly 1350	Asp	Tyr	Ser	Asp	Glu 1355		Asn	Cys	Glu	Asn 1360
	Pro	Thr	Glu	Ala	Pro 1365		Cys	Ser	Arg	Tyr 1370		Gln	Phe	Arg	Cys 1375	-
20	Asn	Gly	His	Cys 1380		Pro	Asn	Arg	Trp 1385		Суз	Asp	Arg	Glu 1390		Asp
	Cys	Gly	Asp 1395		Ser	Asp	Glu	Lys 1400		Сув	Gly	Asp	Ser 1409		Val	Leu
25	Pro	Ser 1410		Thr	Pro	Ala	Pro 1415		Thr	Cys	Leu	Pro 1420		Tyr	Tyr	Arg
	Cys 1425		Gly	Gly	Ala	Cys 1430	Val	Ile	Asp	Thr	Trp 1435		Cys	Asp	Gly	Tyr 1440
30	Arg	Asp	Cys	Ala	Asp 1445		Ser	Asp	Glu	Glu 1450		Сув	Pro	Ser	Leu 1455	
	Asn	Val	Thr	Ala 1460		Ser	Ser	Pro	Ser 1465		Pro	Gly	Arg	Cys 1470		Arg
25	Phe	Glu	Phe 1475		Cys	His	Gln	Pro 1480		Lys	Суз	Ile	Pro 1485		Trp	Arg
35	Arg	Cys 1490		Gly	His	Gln	Asp 1495		Gln	Asp	Gly	Gln 1500		Glu	Ala	Asn
	Cys 1505		Thr	His	Ser	Thr 1510	Leu)	Thr	Cys	Met	Ser 1515	-	Glu	Phe	Lys	Cys 1520
	Glu	Asp	Gly	Glu	Ala 1525		Ile	Val	Leu	Ser 1530		Arg	Cys	Asp	Gly 1535	
	Leu	Asp	Cys	Ser 1540		Glu	Ser	Asp	Glu 1549		Ala	Cys	Ser	Asp 1550		Leu
45	Thr	Val	Tyr 1555	-	Val	Gln	Asn	Leu 1560		Trp	Thr	Ala	Asp 1565		Ser	Gly
	Asn	Val 1570		Leu	Thr	Trp	Met 1575		Pro	Lys	Lys	Met 1580		Ser	Ala	Ala
50	Cys 1585		Tyr	Asn	Val	Tyr 1590	Tyr	Arg	Val	Val	Gly 1595		Ser	Ile	Trp	Lys 1600

	Thr	Leu	Glu	Thr	His 1605	Ser	Asn	Lys	Thr	Asn 161		Val	Leu	Lys	Val 1619	Pen
5	Lys	Pro	Aap	Thr 1620		Tyr	Gln	Val	Lys 1625		Gln	Val	Gln	Cys 1630		Ser
	Lys	Val	His 1635		Thr	Asn	Asp	Phe 1640		Thr	Leu	Arg	Thr 1645		Glu	Gly
10	Leu	Pro 1650	Asp)	Ala	Pro	Gln	Asn 1655		Gln	Leu	Ser	Leu 1660		Gly	Glu	Glu
	Glu 1665	Gly	Val	Ile	Val	Gly 1670		Trp	Ser	Pro	Pro 1675		His	Thr	His	Gly 1680
15	Leu	Ile	Arg	Glu	Tyr 1685		Val	Glu	Tyr	Ser 1690		Ser	Gly	Ser	Lys 1695	
	Trp	Thr	Ser	Glu 1700		Ala	Ala	Ser	Asn 1705		Thr	Glu	Ile	Lys 1710		Leu
20	Leu	Val	Asn 1715		Leu	Tyr	Thr	Val 1720		Val	Ala	Ala	Val 1725		Ser	Arg
	Gly	11e 1730	Gly	Asn	Trp	Ser	Asp 1735		Lys	Ser	Ile	Thr 1740		Val	Lys	Gly
25	Lys 1745		Ile	Pro	Pro	Pro 1750		Ile	His	Ile	Asp 1755		Tyr	qaA	Glu	Asn 1760
	Ser	Leu	Ser	Phe	Thr 1765		Thr	Val	Asp	Gly 1770		Ile	Lys	Val	Asn 1775	Gly
30	Tyr	Val	Val	Asn 1780		Phe	Trp	Ala	Phe 1785		Thr	His	Lys	Gln 1790		Lys
	Lys	Thr	Met 1795		Phe	Gln	Gly	Ser 1800		Val	Ser	His	Lys 1805		Gly	Asn
	Leu	Thr 1810	Ala	Gln	Thr	Ala	Tyr 1815		Ile	Ser	Ala	Trp 1820		Lys	Thr	Asp
35	Leu 1825		Asp	Ser	Pro	Leu 1830		Phe	Glu	His	Val 1835		Thr	Arg	Gly	Val 1840
	Arg	Pro	Pro	Ala	Pro 1845		Leu	Lys	Ala	Arg 1850		Ile	Asn	Gln	Thr 1855	Ala
40	Val	Glu	Cys	Thr 1860		Thr	Gly	Pro	Arg 1865		Val	Val	Tyr	Gly 1870		Phe
	Tyr	Ala	Thr 1875		Phe	Leu	Asp	Leu 1880		Arg	Asn	Pro	Ser 1885		Leu	Thr
45	Thr	Pro 1890		His	Asn	Ala	Thr 1895		Leu	Val	Gly	Lys 1900	-	Glu	Gln	Tyr
	Leu 1905		Leu	Val	Arg	Val 1910		Met	Pro	Tyr	Gln 1915		Pro	Ser	Ser	Asp 1920
50	Tyr	Val	Val	Val	Lys 1925		Ile	Pro	Asp	Ser 1930		Leu	Pro	Pro	Arg 1935	
	Leu	His	Ala	Val	His	Thr	Gly	Lys	Thr	Ser	Ala	Val	Ile	Lys	Trp	Glu

		1940	1943	1950
5	Ser Pro Tyr 195	Asp Ser Pro Asp	Gln Asp Leu Phe Ty 1960	r Ala Ile Ala Val
	Lys Asp Leu 1970	Ile Arg Lys Thr 1975	Asp Arg Ser Tyr Ly	rs Val Lys Ser Arg
10	Asn Ser Thr 1985	Val Glu Tyr Thr 1990	Leu Ser Lys Leu Gl 1995	u Pro Gly Gly Lys 2000
	Tyr His Val	Ile Val Gln Leu 2005	Gly Asn Met Ser Ly 2010	s Asp Ala Ser Val 2015
15	Lys Ile Thr	Thr Val Ser Leu 2020	Ser Ala Pro Asp Al 2025	a Leu Lys Ile Ile 2030
	Thr Glu Asn 2035	Asp His Val Leu	Leu Phe Trp Lys Se 2040	r Leu Ala Leu Lys 2045
00	Glu Lys Tyr 2050	Phe Asn Glu Ser 2055	Arg Gly Tyr Glu Il	e His Met Phe Asp 60
20	Ser Ala Met 2065	Asn Ile Thr Ala 2070	Tyr Leu Gly Asn Th 2075	r Thr Asp Asn Phe 2080
	Phe Lys Ile	Ser Asn Leu Lys 2085	Met Gly His Asn Ty 2090	r Thr Phe Thr Val 2095
25	Gln Ala Arg	Cys Leu Leu Gly 2100	Ser Gln Ile Cys Gl 2105	y Glu Pro Ala Val 2110
	Leu Leu Tyr 2115	Asp Glu Leu Gly	Ser Gly Gly Asp Al 2120	a Ser Ala Met Gln 2125
30	Ala Ala Arg 2130	Ser Thr Asp Val 2135	Ala Ala Val Val Va 21	l Pro Ile Leu Phe 40
· .	Leu Ile Leu 2145	Leu Ser Leu Gly 2150	Val Gly Phe Ala II 2155	e Leu Tyr Thr Lys 2160
35	His Arg Arg	Leu Gln Ser Ser 2165	Phe Thr Ala Phe Al 2170	a Asn Ser His Tyr 2175
	Ser Ser Arg	Leu Gly Ser Ala 2180	Ile Phe Ser Ser Gl 2185	y Asp Asp Leu Gly 2190
40	Glu Asp Asp 2195	Glu Asp Ala Pro	Met Ile Thr Gly Ph 2200	e Ser Asp Asp Val 2205
	Pro Met Val 2210	Ile Ala		
45 (2)	INFORMATION F	OR SEQ ID NO: 3:		
	(A) LEN (B) TYP (C) STR	CHARACTERISTICS GTH: 6961 base p E: nucleic acid ANDEDNESS: doubl	airs	
50		OLOGY: linear TYPE: cDNA to m	RNA	

(ix) FEATURE:
(A) NAME/KEY: sig peptide
(B) LOCATION:178..261

5															. 4		
•		(ix)		TURE													
			(A)) NAI	ME/KI	Σ Υ : π	nat p	sept.	ıae								
			(B)) LO	CATIO	ON:26	52€	816									
										- ·	_						
		(xi)	SEQ	UENC	E DES	SCRII	PTIO	1: S	EQ II	D NO	: 3:						
10															~~~	00100	60
	CCGC	GAGC(CG C	ACAC	GTGA	CGG	CGCC	GCGC	CGC	GCCG	CGC (CGCGG	CCGAC	3C G	GGAC	CCAGC	
	GGCT	GCCC	GG A	GCCC	CGGG	A GC	GCG	CGCG	CGC	GGCC	CCG (3C T		CGGCG	120
	GCGC	GCTG(CAC.	ATTC	TCTC	C TG	3CGG	CGGC	GCC	ACCT	JCA (GCCG	CGTT	JG C	CCGA	ACATG	100
																Met	
										~~~	am a		mma .	OTT N	mmC	ACC 1	228
15	GCG .	ACA (	CGG	AGC .	AGC A	AGG 2	AGG (	GAG	TCG	CGA (	CIC	CCC	TIC (	LIA	Dha	MCC Mb~	220
	Ala	Thr	Arg	Ser	Ser .	Arg A	Arg (	Glu	Ser	Arg	Leu	Pro	Pne .	Leu	Pne	IIII	
				5					10			a. a	ama i	15	1 CC	ccc	276
	CTG	GTC	GCG	CTG	CTG	CCG	CCC	GGG	GCI.	CTC '	TGC	GAG	G1G	TGG	MLG Mh-	722	276
	Leu	Val .	Ala	Leu	Leu	Pro	Pro	GŢĀ	Ala	ren	Cys	GIU	vai	пр	LIIL	ALG	
			20					25			a. a	~~	30	000	mma	CCC	324
20	ACA	CTG	CAC	GGC	GGC	CGC	GCG	CCC	TTA	CCC	CAG	GAG	7	GGC	Dho	7~~	J24
	Thr	Leu	His	Gly	Gly	Arg	Ala	Pro	Leu	Pro	GIN	GIU	AIG	GIY	FIIE	Arg	
		35					40					45	<b>G3.G</b>	acc	aac	ርአጥ	372
	GTG	GTG	CAG	GGC	GAC	CCG	CGC	GAG	CIG	فافانا	CIG	166	CAG	3	Clv	yan	372
	Val	Val	Gln	Gly	Asp	Pro	Arg	Glu	Leu	Arg	Leu	Trp	GIU	Arg	GIY	65	
	50					55					60	cmc	ccc	3.00	707		420
25	GCC	AGG	GGG	GCG	AGC	CGG	GCG	GAC	GAG	AAG	D	CIC	7~~	AGG	Ava	Ara	120
20	Ala	Arg	Gly	Ala	Ser	Arg	Ala	Asp	Glu	ьys	PIO	Leu	AIG	Arg	80	мg	
					70					75	ama	ma c	CCA	CAC		NGC	468
	AGC	GCT	GCC	CTG	CAG	CCC	GAG	CCC	ATC	AAG	GIG	TAC	Clu	CAG	Ual	Ser	100
	Ser	Ala	Ala		Gln	Pro	GIu	Pro	TTE	гÀв	vaı	IYL	GTÅ	95	VAI	261	
				85					90	ama	an.c	mcc	CCC		GNG	λλλ	516
00	CTC	AAT	GAT	TCC	CAC	AAT	CAG	ATG	GTG	GIG	CAC	100	Ala	Cly	Glu	Tare	310
30	Leu	Asn	Asp	Ser	His	Asn	GIn	Met	vaı	vaı	HIS	пр	110	GIY	GIU	цув	
			100					105		~~~	100	OTDC:	110	TTC	CCC	y CC	564
	AGC	AAC	GTG	ATC	GTG	GCC	TTG	GCC	CGG	GAC	AGC	LIG	712	TOU	Δla	Ανα	301
	Ser	Asn	Val	Ile	Val	Ala	Leu	Ala	Arg	Asp	ser	125	AIG	пец	ALU	9	
		115					120	~=~		mam	CAC	T22	CCA	מממ	TCA	דידיר	612
	CCC	AGG	AGC	AGT	GAT	GIG	TAC	GIG	701	IM1	GMC	Time	Clv	Tare	Ser	Phe	
35	Pro	Arg	Ser	Ser	Asp	Val	Tyr	vaı	ser	TAT	140	TAT	GLY	цуз	501	145	
	130					135			mmc	CCC	CCC	CCA	ልልጥ	AAC	ACA		660
	AAT	AAG	ATT	TCA	GAG	AAA	TIG	AAC	Dho	C1.	715	Gly	Aen	Acn	Thr	Glu	
	Asn	Lys	Ile	Ser	Glu	гÀа	Leu	ASI	Pne	155	Ala	Gry	Abii	ruji	160		
					150		m> 0	an a	300	TOO	CCC	GAC	אאכי	ΔΔΔ			708
	GCT	GTG	GTG	GCC	CAG	TTC	TAC	CAC	AGC	CCI	712	AGD	Agn	Taye	Arg	Tvr	
40	Ala	Val	Val		Gln	Pne	lyr	HIS	170	PIO	Ala	Pop	AŞII	175		-1-	
				165		<b></b>	~~~	CAC	170	CTC	TGG	ATC	ΔCG			TTC	756
	ATC	TTC	GCA	GAT	GCC	TAC	GCC	CAG	TAL	LIC	T~~	Tle	Thr	Phe	Asp	TTC	_
	Ile	Phe			Ala	Tyr	ALA	GIII	TAT	beu	TTP	110	190			Phe	
			180			~~~	mmm	185		CCG	ጥጥር	CCC		GCT	GAT	CTC	804
	TGC	AAC	ACC	ATC	CAT	93	TIT	Com	TIO	D×0	Dhe	Ara	Δla	Ala	Asp	CTC	
45	Cys	Asn	Thr	· IIe	HIS	GIY	Pne	Ser	116	FIO	FILE	205	1114		<u>-</u> -	Leu	
		195				~~~	200	330	. Cuter		CTG	CCC	י ייייר	GAC	AGG	TCT	852
	CTA	CTC	CAC	AGT	AAG	33-	Com	AAC	Ten	t.eu	Len	Glv	Phe	Asc	Aro	Ser	
			His	ser	rys	ALA	SeI	Mol	, nea	. Deu	220					225	
	210	·				215	The Co	יא א	• m~	CAT	רבי י	Lutur	. GGC	CAG	ACC	TGG	900
	CAC	: CCC	AAC	AAG	CAG	1.0	TGG	Tare	Cov	. Dan	) Dar	Phe	Glv	Glr	Thr	Trp	
50	His	Pro	Asn	ı r.A.			irp	, Lys	, 361	235	. war				240	)	
50					230	CNC	CTC	. 320	י יירר	لملاسلات د د ت	י י יירעריי	TGG	GGA	ATT		CCC	948
	ATC	ATC	TTA :	CA	GAA	LAC	, UIC	Tare		. Dhe	Ser	Trr	Gly	Ile	Asr	Pro	
	116	e Met	: 116			HIS	val	. uya	250					255	5	_	
				245	•				230	•							

55

			2.6	Ω		C ACC		- 191	- 11	E (-1	u Ar	g ni	B GT	u Pı	o s	er	Glv	996
5	TA	C TC	C AC	G GT	ጥ ጥጥ	CGA Arg	AGT Ser	ACA Thi		- mm	~	C CA	27 G TC n Se	0			_	1044
	CA G1: 29	G GA n Gl 0	A GT u Va	G ATO	C TTO	GAG Glu	GAA Glu	GTG Val	AG/	A GA	C TT	I CAG	s G CT n Le	T CG	G G	AC sp	AAG Lys	1092
10	TAC Ty:	C AT	G TT	r GC	T ACA Thr	AAG Lys	GTG Val	GTG Val	CAT His	CTC	TT(	G GGG	C AG	T CC r Pr	A C	rg eu	305 CAG Gln	1140
				325	G CTC	TGG Trp	VAI	Ser	220	GGC Gly	CGC Arg	l rat	3 Pro	o Me	t Aı	g	Ala	1188
15			340	)		AGA Arg	1113	345	Ile	AAC Asr	1 GIV	тул	Ty	r Il	e Al	.a	Asp	1236
		355	5			GTG Val	360	GTG Val	TGI Cys	vaı	. Ser	His	AGC Sea	AA As	n As	n	Arg	1284
20	370				110	TCG Ser 375	GIU	AIA	GIU	СТУ	Leu	AAC Lys	TTO Phe	Se:	r Le	u	Ser	1332
					390	TAC Tyr	TYL	1111	PLO	GTA	GGG	GCC Ala	Gl _}	Se:	r As	p	Thr	1380
25			9	405		GCA Ala	ASII	GIU	430	TTT Phe	GCT Ala	Asp	Phe	His	3 Ar	g `	Val	1428
		•	420			GTC Val	TYL	425	мта	Inr	Leu	ile	Asn	GG;	TC Se	r I	Met	1476
30		435			1100	AGA Arg	440	vai	TTE	Thr	Phe	Asp	AAA Lys	GG( Gl _y	Gl;	Y ?	Thr	1524
	450	-		Deu	GIII	GCT Ala 455	PIO	Ата	Pne	Thr	GGG Gly	TAT Tyr	Gly	Glu	Ly	3 :	Ile	1572
35		-70		Deu	470	GAG Glu	GIĀ	Cys	ser	Leu	His	Leu	Ala	Gln	Arg	g (	CTC Leu	1620
				485	11011	CTC Leu	GIII	ren	490	Arg	Met	Pro	Ile	Leu	TC(	: I	уs	1668
<b>4</b> 0			500	110	Gry	CTC . Leu	116	116	ALA	Thr	GIY	Ser	Val	GGA Gly	AA(	A	lsn	1716
		515		-,-		AAC Asn	520	TAT	TTE	ser	ser	Ser	Ala	Gly	Ala	ı A	rg	1764
45	<b>5</b> 30				<b></b>	CCT ( Pro (	JIY	PIO	nis	ıyr	Tyr	Thr	Trp	Gly	Asp	H	lis	1812
	-				550	GCC Ala	LIE .	nia '	GIN	GIY	mec	Glu	Thr	Asn	Glu	L	eu	1860
50	•	- <b>/ -</b>		565	ASII '	GAA ( Glu (	er A	GIU :	I'UL 570	Trp	Lys	Ala	Phe	Thr	TTC Phe	T	er	1908
		-,,	580	V41	FIIE	GTG 1	yr	517 I	Leu	Leu	Thr	Glu	Pro	Gly	Glu	L	уs	1956
		-100	GIC	110	ACC A	ATC 1	TT (	∍GC :	rcc .	AAC	AAG	GAG	AAC	GTG	CAC	A	GC	2004

	Ser		Val	Phe	Thr		Phe	Gly	Ser	Asr.	Lys	Glu . 605	Asn '	Val	His	ser	
	TGG	595 CTC	ATC	CTC	CAG Gln	CTC	ገል ፈ	GCC Ala	ACA Thr	GAC Asp	GCC Ala	CTG Leu	GGG Gly	GTT Val	CC <b>T</b> Pro	TGC Cys	2052
5	C10					615					620					625	
	A CA	GAG	AAC	GAC	TAC	AAG	CTC	TGG	TCA	CCA	TCT	GAT	GAG	CGG	GGG	AAT	2100
	Thr	Glu	Asn	Asp		ràs	Leu	Trp	ser	635	Ser	Asp	GIU	Arg	640	ASII	
	CAG	тст	TTC	ידידיי	GGA	CAC	AAG	ACT	GTT	TTC	AAA	CGG	AGG	ACC	CCG	CAC	2148
	Glu	Cvs	Leu	Leu	Gly	His	Lys	Thr	Val	Phe	Lys	Arg	Arg	Thr	Pro	His	
10		_		CAE					650					655			2196
	GCC	ACA	TGC	TTT	AAC	GGA	GAA	GAC	TTT	GAC	AGG	CCG Pro	Val	Val	Val	Ser	2130
			$\epsilon \epsilon \Delta$					665					6/0				
	AAC	TGC	TCC	TGC	ACC	CGG	GAG	GAC	TAT	GAG	TGT	GAC	TTT	GGC	TTC	CGG	2244
	Asn	Cys	Ser	Суз	Thr	Arg	Glu	Asp	Tyr	Glu	Cys	Asp	Phe	GIA	Phe	Arg	
15		C75					680					600					2292
	ATG	AGT	GAA	GAC	TTG	GCA	Leu	GAG	Val	Cvs	Val	CCA Pro	Asp	Pro	Gly	Phe	
						495					700					, 0 5	
	~~~	GGA	AAG	TCC	TCC	CCT	CCA	GTG	CCT	TGT	CCC	GTG	GGC	TCT	ACG	TAC	2340
	Ser	Gly	Lys	Ser	Ser	Pro	Pro	Val	Pro	Cys 715	Pro	Val	GIA	Ser	720	TYL	
20		~~~	mas.	303	710	ጥአር	CGG	λAG	ΔТТ	TCT	GGG	GAC	ACC	TGT		GGA	2388
	AGG	Ara	Ser	Arg	Glv	Tyr	Arg	Lys	Ile	Ser	Gly	Asp	Thr	Cys	Ser	Gly	
	_			775					7.30					133			2426
	GGA	GAT	GTT	GAG	GCA	CGG	CTA	GAA	GGA	GAG	CTG	GTC	Pro	TGT	Pro	Leu	2436
	Gly	Asp			Ala	Arg	Leu	745	GIY	GIU	Leu	Val	750	Cys	rio	D Cu	
25	CCN	CAA	740 GAG	አአር	GAG	TTC	ATC	CTG	TAC	GCC	: ACG	CGC	AAG	TCC	ATC	CAC	2484
	Ala	Glu	Glu	Asn	Glu	Phe	Ile	Leu	Tyr	Ala	Thr	Arg	Lys	Ser	Ile	His	
							760					/00					2532
	CGC	TAT	GAC	CTG	GCT	TCC	GGA	ACC	Thr	GAU	Glr	TTG Leu	Pro	Leu	Thr	Gly	2332
	770					775					701	,				, 0 3	
30	770 TTG		GCA	GCA	GTG	CCC	CTG	GAC	TTT	'GA	TAT	GAG	CAC	AAC	TGC	CTG	2580
	Leu	Arc	Ala	Ala	Val	Ala	Leu	Asp	Phe	AS	э туг	Glu	His	Asn	Cys 800	пеп	
					700					/4.	•				000	AAC	2628
	TAT	TGC	TCT	GAC	CIG	Δla	Len	Asc	Val	Ile	e Gli	Arg	Leu	Cys	Lev	Asn	
				0 0 5					810	1				01-	,		
35	GGC	AG	ACA		C 2 2	GAG	GTG	ATC	ATO	AA	C TC	r GAC	CTG	GAC	ACC	GTA	2676
	Gly	/ Sei	Thi	c Gly	/ Glr	ı Glu	Val	. 11e	; ITE	AS	n Se	r Asp	ь Leu 830	Git	1 1111	. Val	
	~ .		820) זרטיי	is ubstract	ር ረን ን	ccc	825 יייטיי	'AGC	CA.	A TT	A CTT			GTO	GAC	2724
	GAZ	ı Ala	i iio	ı Ala	a Phe	Glu	Pro	Lev	Sei	c Gl	n Le	ı Lev	ı Tyr	Tr	va.	L Asp	
		0.01	_				947	3				043	,				2772
40	GC	A GG	C TT	r aa	A AAC	ATC	GAC	GTA	A GC	C AA	n Dr	A GAT	CIS C	, GAV	o Phe	CGA Arg	2,,2
																e Arg 865	
	85°		CGT	C GT	C AA	י יייר	TO	GTO	G CT	G GA	T CG	G CCC	CGG	GC	CCT	G GTC	2820
	Le	u Th	r Va	l Va	l Ası	n Sei	: Sei	r Vai	l Le	u no	א ע	g Pro	o Arg	Al			
					07	_				~ /						_	2868
45	CT	T GT	G CC	C CA	A GA	A GGC	3 ATC	. ATC	اتات Ph	e Tr	n Th	r Ası	o Tri	G1	y As	C CTG p Leu	
				00	_				ж ч	()					_		
	AA	G CC	T GG	~ 20	T T T T	T CG	G AG	CAA	C AT	G GA	C GG	A TC	r gc	C GC	C TA	T CGC	2916
	Lу	s Pr	o Gl	y Il	е Ту	r Ar	g Se	r As	n Me	t As	p Gl	y Se	r Ala 91	2 11	а ту	r Arg	
			~ ~	^				40	_				2 1	•		C GAT	
50	CT T.e	C GI	G TC	GAن 1. صرا	لفان. عΣ ۱۱	n Va	J AA	s Tr	p Pr	o As	n Gl	y Il	e Se	r Va	l As	p Asp	
		~ 4					כים	n .				74					
	CA	a ma		C TA	C TG	G AC	G GA	T GC	C TA	C C	rg ga	C TG	CAT	I GA	.G CG	C ATC	3012
	Gl	n Tr	p Il	e Ty	r Tr	p Th	r As	p Al	a Ty	rr Lie	eu As	P CA	9 II	- GI	u Al	g Ile	•

	930 ACG TT	C AGC (GC CAG	935 CAG	CGC TCC	GTC	AIC	940 CTG	GAC	ACA	CTC	: cce	945 CAC	3060
5			950		g Ser	. vaı	116	ren	Asp	Arg	Leu	Pro	His	2060
	_	T GCC A	65		гие пуз	ASII	GIU	тте	Tyr	Trp	Asp	Asp	Trp	3108
10		G CTC A n Leu S 980			ny Ala	Ser	гуя	ıyr	Ser	Gly	Ser	Gln	Met	3156
	99	T CTG C e Leu A 5		1	CTC ACG	GGG Gly	Leu	met	GAC Asp	Met	Lys	Ile	Phe	3204
	TAC AA	G GGG A	AG AAC	ACA C	GA ACC	አአጥ	CCC	TGT Cys	1005 GTA Val					3252
15	AGC CT	G CTG T Leu C	GC CTG ys Leu	CCC A	GA GCC	220	220	1020					1025	3300
	CCA GA	GGC G	TG GCC	AGC A	CT CTC	CTC	1032	maa .	~~~			104)	3348
20	GAC TG	CCTA	AG GGC	TAC G	AG CTG	770	, አአሮ	A A C :	3.CC 1	nom	1055	5	_	3396
	GAA GAG	1060 ACC TO	T CTG	CGC A	106!	Dys 5	asn	Asn :	Inr (Cys 1070	Val	Lys	Glu	
<i>25</i>	107	'5 'AGC A'	re tree	TGG T	OSO CAT	Tyr	Arg	Cys S	Ser 1	Asn (Gly	Asn	Cys	3444
23	ATC AAC Ile Asn 1090		-C 11p	1095	ys Asp	Pne .	Asp .	ASD A	Asp (ys (Gly	Asp	Met	3492
	AGC GAC Ser Asp	,	1110	Cys F.	IO IIII	Inr	11e (Cys A	rab I	eu 1	Asp	Thr	CAG Gln	3540
30	TTC CGT Phe Arg	11	.25	SEI G.	ry inr	TGC :	ATC (Pro I	eu S	er 1	ľyr	AAA Lys	TG T Cys	3588
	GAC CTC Asp Leu	1140	Pp	cys G.	1145	AAC A	ser A	Asp G	ilu A	GG (łis	TGT Cys	Glu	3636
35	ATG CAC Met His 115	5	J 1119	11	AC GAA sp Glu	TAC I	Asn (ys S	GC Ter S	CG C	ly	Met	Cys	3684
	ATC CGC Ile Arg 1170			1175	a Asp	GIY A	asp A	AAC G	AC T sp C	ys A	urg .	Asp '	Trp	3732
40	TCC GAC Ser Asp		1190	cys II	II AIA	TIE 1	TAT C	AC A lis T	hr C	ys G	lu i	Ala :	Ser	3780
	AAC TTC Asn Phe	12	C CGC 1 s Arg 1	01	y nis	TGC A	TC C	TO G	Tu A	rg T	Lb 1	Ala (Cys	3828
4 5	GAC GGC Asp Gly	GAC GC Asp Al 1220	C GAC :	rgc ca	C CAT /	~~~ =	CT G er A	AT G	Lu A	AT C sp P	215 CA (ro !	GCC A	AAC Asn	3876
	TGT GAG Cys Glu 1239	AAG AA	G TGC A	AAC GG Asn Gl	1223 C TTC (ode m	·CC	CG A	AC G sn G	230				3924
50	CCC TCC Pro Ser 1250	ACC AA Thr Ly	G CAC T S His C	GT GA	C GGC (p Gly I	CTG C Leu H	AC G	AT TO	245 GC T(ys Se	CG G	AC (GC T	CCC Ser	3972
	GAC GAG Asp Glu	CAG CA	C TGC G	AG CC	C CTG 7	-Aa I	CAC	260 GG TT rg Ph	rc An ne Me	rg ga	sp F	1 TC 6 he V .280	.265 TG 'al	4020

	TGT :	AAG Lys	AAC Asn	CGC Arg 1285	Gln	CAG Gln	TGC Cys	CTC Leu	TTC Phe 1290	his	TCC Ser	ATG Met	٧۵٦	TGC Cys 1295	Asp	GGG Gly	4063
5	ATC .	ATC Ile	Gln	TGC Cys	CCT	GAC Asp	GGC Gly	TCC Ser 1305	GAC Asp	GAG	GAC Asp	CCA Pro	GCC Ala 1310	Phe	GCA Ala	GGA Gly	4116
		Ser	Arg	CAC	CCC Pro	GAG Glu	TTC Phe 1320	CAC His	AAG	GTG Val	TGC Cys	GAT Asp 1325	GAG Glu	TTC	GGC Gly	TTC Phe	4164
10	C7 C	1315 TGT Cvs	CAC	AAC Asn	GGC Gly	GTG Val	TGC	ATC	AGC Ser	TTG Leu	ATC Ile	TGG	AAG	TGC Cys	GAC Asp	GIY	4212
	1330	G 3 TF	CAC	TOC	GGG Gly	1335	ፐልሮ	TCC	GAC	GAG	GCC	AAC	TGT	GAA	AAC	CCC	4260
		C 3 3	ccc	ccc	1350) ייניר	TCC	CGC	TAC	TTC	CAG	TTC	CGG	TGT	GAC	AAT	4308
15	Thr	Glu	Ala	Pro	Asn	Cys	Ser	Arg	Tyr 137	Pne 0	GIN	Pne	Arg	1379	wab 2	ASII	4356
	Gly	His	Cys	Ile	CCC	Asn	Arg	Trp	Lys 5	Cys	Asp	Arg	1390	ASII	Asp	Cys	
20	GGG Gly	GAC Asp	TCC	TCC	GAC Asp	GAG Glu	Lys	Asp	TGT Cys	GGA Gly	GAT Asp	TCA Ser 140	His	GTA Val	CTT Leu	CCG Pro	4404
	TCT	139! ACG	A CT	CCT Pro	GCA Ala	CCC Pro	TCC Ser	ACG	TGT Cys	CTG Leu	CCC Pro	AAT	TAC	TAC Tyr	CGC Arg	TGC Cys	4452
	1410	222	acc	ccc	ጥርር	141!	5 ATA	GAC	ACG	TGG	142 GTT	o TGT	GAC	GGG	TAC	1425 CGA Arg	4500
25		maa	CCN	CNT	143	0 יייכיר	GAC	GAG	GAA	GCC	TGC	ccc	TCG	CTC	CCC	AAT	4548
	Asp	Cys	Ala	Asp	Gly	Ser	Asp	Glu	Glu 145	Ala O	Cys	Pro	ser	145	5	Asn	4596
30	Val	Thr	Ala	Thr	Ser	Ser	Pro	Ser	GIn	Pro	GLY	Arg	147	o O	MIG	FIIC	4544
	GAG Glu	Phe	Glu	TGC	CAC His	CAG Gln	CCA Pro) Lys	Lys	G TGC	: ATC	CCT Pro	ASII	Trp	Ace	CGC Arg	4644
	TGT Cvs	177 GAC Asp	~~~	CAT	CAG Gln	GAT Asp	TGC	CAG	GAT Asp	GGC Gly	/ Gli	GAC Asp	GAG	GCC	AA(TGC Cys	4692
35	149	0		, ACC	י אככ	149 TTC	5 . acc	ጉ ጥርር	י אדנ	G AGO	150 TG0	GAC	TTC	AAG	TG	1505 GAG Glu	4740
	G 3 M				151 TGC	.0 י אידר	י פידיני	ב כידו	3 TC	A GAZ	LO A CGO	TG	GAC	GGG	TT	CTG	4788
40	Asp	Gly	Glu	152	i Cys 25 r cac	i Ile	· Val	L Lev	150 151 3 AAG	G GC(G GC(Arg	G AG	r GAT	153 GAC	35 3 TT	A ACT	4836
	Asp	Суя	Sei	. Ası	p Glu	ı Sei	: Ası	9 GI1	ц Ly : 45	S Ala	а су	s 5e.	159	50 50	ם חכ	u 1111	4884
	GTA Val	. Ty	r Lys	A GT	A CAC	AA!	CT. Lev 15	u GL	G TG	G AC	A GC r Ala	r GA a As 15	b bue	e Se	r Gl	G AAT y Asn	4004
45	GT(Val	159 AC: L Thi	T 7777	G AC	C TG(p Met	G CGG	G CC	C AA o Ly	A AA s Ly	s me	G CC	C TC	r GC r Al	r GC a Al	T TGT a Cys 158!	4932
	157	70 N. 117 A	C 33	c cm	ር ጥል	15' ימיזי רי	75 T AG	A GT	т ст	T GG	A GA	G AG	C AT	A TG	G AA	G ACT	4980
50	CMI)	- CN	~ .~	ጥ ሮአ	15: C AG	90 C DD'	44 T	G AC	A AA	C AC	yo TGT	A TT	A AA	A GT	G TT	GAAA	502
				16	ΛE				1.6	110				10	_	u Lys C AAG	507

	Pro	As	o Th 16	r Th 20	r Ty	r Gl	n Va	1 Ly 16	s Va	l Gi	.n Va	1 G1	n Çy	∕e I.e	n S	er Lys	
	GTG	CA	CAA	C AC	C AA	T GA	С Т Т	ጥ ርጥ	ວັນຕ	a			_ Te	30			•
5	Val	Hi	aA e	n Th	r As	n Ası	p Ph	e Va	1 Th	~ T.a	U AG	A AL	ar. CC	A GA	G GC	A TTG y Leu	5124
5		163	3 5				16	40		. De	u AL	9 10	r Pr	O G1	u Gl	y Leu	
	CCA	GA	GC	c cc	T CA	J AA	س.	C CN	c cm	G T/C	c cm		45		_	G GAA	
	Pro	Asp	Al	a Pr	o Gli) As	n T.e.	n Gl	n Ta		- CI	C CA	C GG	G GA	A GA	G GAA u Glu	5172
	165	0				16	55	u (31)	n ne	u se	r re	u Hi	s GI	y Gl	u Gl	u Glu	
	GGT	GTG	TA	T GT	G GGG	רא ר	TC(G 7/C		F 00	16			_		1669 C CTC	5
	Gly	Val	Il	e Va	1 Gly	, His	z Tr	n Se	c (c.,	1 CC	C AC	C CA	CAC	C CA	C GG	C CTC y Leu	5220
10	•				16	70	J 11	o se	L PIC) PI	o in	r Hi	s Th	r Hi	s Gl	y Leu	
	ATT	CGC	GA	A TA	CAT	CT2	CAC	י מידי	r 200	10	/3 ~ 3 ~				16	80	
	Ile	Ara	Gli	ı Tv	r Ile	Val	GI	J The	. AGC	- AG	G AG	I. GG.	r rc	CAA	G GT	G TGG l Trp	5268
				16			. GI	ı ıyı	. sei	Ar	g se	r GI	y Se:	r Ly:	s Va	l Trp	
	ACT	TCA	GAZ	A AG	GGCT	י הכיו	י אכיי	ר אאר	169	7 U				16	95	G TTG	
	Thr	Ser	Gli	1 Arc	T Ala	Δ1:	Ser	r AA	. Db.	AC.	A GA	A AT	A AA	G AA	C TT	G TTG	5316
15			170	00	J	• ••••	. 561	17/	1 PITE	: 111	r GII	1 II	∋ Ly	s Ası	ı Le	G TTG u Leu	
	GTC	AAC	ACC	CTC	3 TAC	' ACC	CTC	170 2 AG2)				17:	10		T GGG	
	Val	Asn	The	Lei	ı Tvr	The	· Val	A TO	* W-1	7.7	r GCC	GT(3 AC	G AG	r cg	T GGG	5364
		171	5			* ***	172	י עדה	y vai	. Ala	a Ala	a va.	LTh	r Sei	r Ar	T GGG g Gly	
	ATA	GGA	AAC	TGO	G AGC	GAT	TCC	יט מממי	TOO			172	25 			A AAA	
	Ile	Glv	Asr	Trr	Ser	Acr	Sor	· Dane	Com	ATT	ACC	ACC	GTO	3 AA	A GG	A AAA	5412
20	1730) 1			,	173	, Set	пув	ser	116	e Thi	Thi	· Val	l Lya	Gl:	A AAA y Lys	
20	GCG	ATC	CCC	CCZ		ል አጥ	ັ້ນກາດ		, y.com		174	10				1745 F TCC	i
	Ala	Ile	Pro	Pro	Pro	Agn	Tle	. Uic	. Tlo	GAC	AAC	TAC	GAT	C GAZ	AA	TCC	5460
																r TCC n Ser	
	CTG	AGT	TTT	ACC	CTG	ACC	СТС	CAT	ccc	175					176		
	Leu	Ser	Phe	Thr	Leu	Thr	Val	Acu	. Cl.	AAC	ATC	AAC	GTC	AA1	GG	TAT Tyr	5508
				176	5		val	лър	177	ASD	1 TTE	: гуз	val			/ Tyr	
<i>25</i>	GTG	GTG	AAC	CTT	TTC	TGG	GCA	طبطين	CAC	` 3 .CC				177	' 5.	AAA F	
	Val	Val	Asn	Leu	Phe	Trn	Ala	Dhe	yer	The	CAC	AAA	CAA	GAG	AAC	AAA Lys	5556
			178	0				178	e E	THE	nis	гув	Gin	GIu	Lys	Lys	
	ACC	ATG	AAC	TTC	CAA	GGG	AGC	TCA	CTC	TCC			179			CTG	
	Thr	Met	Asn	Phe	Gln	Glv	Ser	Ser	Val	200	CAC	AAA	GTT	GGC	'AA'	CTG Leu	5604
		1795	5			1	180	ດີ້	141	SET	птэ	Lys	_var	GLY	Asr	Leu	
30	ACA	GCA	CAG	ACG	GCC	TAT	GAG	እ ጥጥ	TCC	GCC	TCC	180				TTG	
	Thr	Ala	Gln	Thr	Ala	Tvr	Glu	Tle	Ser	Ala	TV	31-	AAG	ACT	GAC	TTG Leu	5652
	GGC	GAT	AGT	CCT	CTG	TCA	Tur	GAG	САТ	GTC	3.00	3.00	202	000		1825 CGC	
	Gly .	Asp	Ser	Pro	Leu	Ser	Phe	Glu	His	Val	The	Th~	AGA	- GGG	GTT	CGC Arg	5700
35	CCA	CCT	GCT	CCT	AGC	CTC	AAG	GCC	AGG	CCT	_ እጥ/ጋ	አአጥ	CNC	3.00	184	_	
	Pro 1	Pro	Ala	Pro	Ser	Leu	Lys	Ala	Arg	Ala	Tle	yen	CAG	The	GCA	GTG	5748
	GAA :	IGC	ACC	TGG	ACA	GGC	CCC	AGG	አአጥ	CTC	GTG	тат	GGC	185!	_	m. m	
	Glu (-			Thr	Gly	Pro	Arq	Asn	Val	Val	Tvr	Glv	Tla	Dho	TAT	5796
40	GCC A	ACA	TCC	TTC	CTG	GAC	CTC	TAC	CGC	AAC	CCA	AGC	300		N.C.C	3.00	5044
				Phe	Leu	Asp	Leu	Tyr	Arq	Asn	Pro	Ser	Ser	Lau	Th~	ACG	5844
	CCG (CTG	CAC	AAC	GCA	ACC	GTG	CTC	GTC	GGT	AAG	~~~	~-~	CAG	TAT	CTC	E000
		Leu	His	Asn	Ala	Thr	Val	Leu	Val	Gly	Lvs	Asp	Glu	Gla	Trr	LOU	5892
45	TTT (TG	GTC	CGG	GTG	GTG	ATG	CCC	TAC	CAA	000	~~~	TCC	TCG	CAC	_	E040
	Phe I	-eu	Val	Arg	Val	Val	Met	Pro	Tyr	Gln	Glv	Pro	Ser	Ser	Acn	Ture	5940
	OTT -				1910		*		•	1915	5			001	1920	, I Y L	
	GTG G	FIC (GTG	AAG	ATG	ATC	CCG	GAC	AGC	AGG	CTT	CCT	CCC	CGG			E000
	Val V	al '	Val	Lys	Met	Ile	Pro	Asp	Ser .	Ara	Leu	Pro	Pro	Ara	Lic	Tan	5988
	Cam -			1925	5			-	1930	-				1935		₽e a	
50	CAT G	CC (STT	CAC	ACC	GGC .	AAG	ACC	TCG .	CCC	GTC	ATC	AAG			TCG	6026
30	His A				Thr	Gly	Lys	THE	ser.	Ala	Val	Ile	Lvs	Tro	Glu	Ser	6036
	CC0 ~		1940					1945					1950			Jer	
	CCC T	AC (JAC	TCT	CCT	GAC	CAG	GAC	CTG '	TTC	TAT	GCG	> ~~		GTT	ΔΔΔ	6084
	Pro T	AL I	rsb	ser	Pro .	Asp	Gln .	Asp	Leu	Phe	Tyr	Ala	Ile	Ala	Val	Lvs	0004
											-					-,-	

		1955					1960)				1965					
	GAT	CTG	ATA	CGA	AAG	ACG	GAC	CGG	AGC	1·AC	ΛΑΛ	STC	AAG	TCC	CGC	AAC	6132
			Ile	Arg	Lys	Thr	Asp	Arg	Ser	Tyr	Lys	Val	Lys	Ser	Arg	Asn	
5	1970			~~~		1975	ama	100		CTT C	1980		CCA	ccc		1985	6180
	AGC	ACC	GIG	GAG	TAC	ACC	CIG	AGC Ser	AAG	LAU	Glu	Pro	Clv	GUU	Tara	Tur	9160
	ser	Thr	vaı	GIU	1990		Leu	Ser	гув	1999		FIO	GIY	Gry	2000		
	CAC	GTC	ΑΤΤ	GTG	CAG	CTG	GGG	AAC	ATG			GAT	GCC	AGT			6228
	His	Val	Ile	Val	Gln	Leu	Gly	Asn	Met	Ser	Lys	Asp	Ala	Ser	Val	Lys	
				2005	;				2010)				2019	5		
10	ATC	ACC	ACC	GTT	TCG	TTA	TCG	GCA	CCC	GAT	GCC	TTA	AAA	ATC	ATA	ACA	6276
	Ile	Thr			Ser	Leu	Ser	Ala	Pro	Asp	Ala	Leu	Lys	, i i e	TTe	Thr	
			2020)	ama		omo.	2025 TTC	TCC	***	አርማ	CTA	2030		AAG	GAA	6324
	GAA	AAT	GAC	CAC	GTC	CTT	Lau	Phe	TGG	Lve	Ser	LEN	Δla	Len	Lvs	Glu	0324
	GIu	2039		HIS	vaı	reu	2040		пр	шув	Ser	2045	71a	<u> L</u> Cu	2,0	014	
15	AAG	ТΔТ	بلملىك	AAC	GAA	AGC	AGG	GGC	TAC	GAG	ATA	CAC	ATG	TTT	GAT	AGC	6372
	Lvs	Tvr	Phe	Asn	Glu	Ser	Arg	Gly	Tyr	Glu	Ile	His	Met	Phe	Asp	Ser	
	2050	١ -				2055	5				2060)				2065	
	GCC	ATG	AAT	ATC	ACC	GCA	TAC	CTT	GGG	AAT	ACT	ACT	GAC	AAT	TTC	TTT	6420
	Ala	Met	Asn	Ile			Tyr	Leu	GIA			Thr	Asp	Asn	208	Pne	
					2070) 120	3 m/c	CCT	CNC	207	ጋ ጥልሮ	A C'A	ጥጥር	ACG		CAG	6468
20	AAA	ATT	TCC	AAC	CTG	AAG	Mat	GGT	Uic	Ven	Tyr	Thr	Phe	Thr	Val	Gln	0400
	га	TIE	Ser	208		пур	MEC	GIY	209	0		****		209	5		
	GCA	CGA	TGC	CTT	TTG	GGC	AGC	CAG	ATC	TGC	GGG	GAG	CCT	GCC	GTG	CTA	6516
	Ala	Arg	Cys	Leu	Leu	Gly	Ser	Gln	Ile	Cys	Gly	Glu	Pro	Ala	Val	Leu	
			210	a				210	5				211	U			CE C 4
25	CTG	TAT	GAT	GAG	CTG	GGG	TCT	GGT	GGC	GAT	GCG	TCG	GCG	ATG	CAG	GCT	6564
	Leu			Glu	Leu	GIĀ	Ser 212	GIA	GIY	Asp	Ala	212	Z WIG	Mec	GIII	Ala	
	222	211	יייטייי ב	አ ርጥ	ርልጥ	GTC	GCC	ᢅᢎᡳᠸ	GTG	GTG	GTG			CTG	TTT	CTG	6612
	Ala	AGG	Ser	Thr	Asn	Val	Ala	Ala	Val	Val	Val	Pro	Ile	Leu	Phè	Leu	
	212	۸				213	5				214	0				2145	
	ΔΤΔ	CTG	CTG	AGC	CTG	GGG	GTC	GGG	TTT	GCC	ATC	CTG	TAC	ACG	AAG	CAT	6660
30	Ile	Leu	Leu	Ser			Val	Gly	Phe	Ala	_Ile	Leu	Tyr	Thr	Lys	His	
					215	0				215	5		3.00	C3.C	216		6708
	CGG	AGG	CTG	CAG	AGC	AGC	TTC	ACC	GCC	TTC	31a	AAC	AGC	UAC.	Trac	AGC	6708
	Arg	Arg	Leu	G1n 216		ser	Pne	inr	217	n n	Ala	ASII	361	217	'5	Ser	
	TOC	303	CTC	210	ישרר ס	GCC	אידר	י יייר	TCC	TCT	· GGG	GAT	GAC			GAG	6756
35	Ser	ACA	Leu	Glv	Ser	Ala	Ile	Phe	Ser	Ser	Gly	Asp	Asp	Leu	Gly	glu glu	
			218	٥				218	5				219	0			
	GAT	GAT	CAA	GAT	GCT	CCT	ATG	ATC	ACI	GGA	TTT	TCG	GAC	GAC	GTC	CCC	6804
	Asp	Asp	Glu	Asp	Ala	Pro	Met	: Ile	Thr	Gly	, Phe	Ser	Asr	Asp	Val	Pro	
	_	219	5				220	0				220		יוויריותי ז	ייי א א		6856
	ATG	GTG	ATA	GCC	TGA	AAGA	GCT	TTCC	TCAC	TA	JAAAC	CAAA	IT GO	1611	A-A-V T I	•	6656
40			Ile	Ala	Ļ												
	221	. U זייט אַרדא	ማር እ	מממידי	ርልጥል	ייים.	יתאדים	יייייני	וידי בי	TTA	AAGA	TGC	ACT	TGA	GTT	CAATA	6916
	ርጥጥ T T.T	יירא די די איז	TUPA ATTE	TATO	GGCC	'AA	AAC	AAAA	C A	VAAA	AAAA	AAA	AA				6961
	GII		* * * **														
	(0)			MITON	T POE	000	מד ר	MO.	Δ.								

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 300 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA to mRNA

55

50

	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:	
_	ATATCCACAT TGACAGCTAT GGTGAAAATT ATCTAAGCTT CACCCTGACC ATGGAGAGTG	60
5	ATATCAAGGT GAATGGCT.AT GTGGTGAACC TTTTCTGGGC ATTTGACACC CACAAGCAAG	120
	AGAGGAGAAC TTTGAACTTC CGAGGAAGCA TATTGTCACA CAAAGTTGGC AATCTGACAG	180
	CTCATACATC CTATGAGATT TCTGCCTGGG CCAAGACTGA CTTGGGGGAT AGCCCTCTGG	240
10	CATTTGAGCA TGTTATGACC AGAGGGGTTC GCCCACCTGC ACCTAGCCTC AAGGCCAAAG	300
	(2) INFORMATION FOR SEQ ID NO: 5:	
	(i) SEQUENCE CHARACTERISTICS:	
15	(A) LENGTH: 6642 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
20	(ii) MOLECULE TYPE: cDNA to mRNA	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:	
25	ATGGCGACAC GGAGCAGCAG GAGGGAGTCG CGACTCCCGT TCCTATTCAC CCTGGTCGCA	60
	CTGCTGCCGC CCGGAGCTCT CTGCGAAGTC TGGACGCAGA GGCTGCACGG CGGCAGCGCG	120
	CCCTTGCCCC AGGACCGGGG CTTCCTCGTG GTGCAGGGCG ACCCGCGCGA GCTGCGGCTG	180
30	TGGGCGCGC GGGATGCCAG GGGGGCGAGC CGCGCGGACG AGAAGCCGCT CCGGAGGAAA	240
	CGGAGCGCTG CCCTGCAGCC CGAGCCCATC AAGGTGTACG GACAGGTTAG TCTGAATGAT	300
	TCCCACAATC AGATGGTGGT GCACTGGGCT GGAGAGAAAA GCAACGTGAT CGTGGCCTTG	360
5	GCCCGAGATA GCCTGGCATT GGCGAGGCCC AAGAGCAGTG ATGTGTACGT GTCTTACGAC	420
	TATGGAAAAT CATTCAAGAA AATTTCAGAC AAGTTAAACT TTGGCTTGGG AAATAGGAGT	480
	GAAGCTGTTA TCGCCCAGTT CTACCACAGC CCTGCGGACA ACAAGCGGTA CATCTTTGCA	540
0	GACGCTTATG CCCAGTACCT CTGGATCACG TTTGACTTCT GCAACACTCT TCAAGGCTTT	600
	TCCATCCCAT TTCGGGCAGC TGATCTCCTC CTACACAGTA AGGCCTCCAA CCTTCTCTTG	660
	GGCTTTGACA GGTCCCACCC CAACAAGCAG CTGTGGAAGT CAGATGACTT TGGCCAGACC	720
5	TGGATCATGA TTCAGGAACA TGTCAAGTCC TTTTCTTGGG GAATTGATCC CTATGACAAA	780
	CCAAATACCA TCTACATTGA ACGACACGAA CCCTCTGGCT ACTCCACTGT CTTCCGAAGT	840
	ACAGATTTCT TCCAGTCCCG GGAAAACCAG GAAGTGATCC TTGAGGAAGT GAGAGATTTT	900
,	CAGCTTCGGG ACAAGTACAT GTTTGCTACA AAGGTGGTGC ATCTCTTGGG CAGTGAACAG	960
,	CAGTCTTCTG TCCAGCTCTG GGTCTCCTTT GGCCGGAAGC CCATGAGAGC AGCCCAGTTT	1020
	GTCACAAGAC ATCCTATTAA TGAATATTAC ATCGCAGATG CCTCCGAGGA CCAGGTGTTT	1080

	GTGTGTGTCA	GCCACAGTAA	CAACCGCACC	AATTTATACA	TCTCAGAGGC	AGA(GGGGCTG	1140
	AAGTTCTCCC	TGTCCTTGGA	GAACGTGCTC	TATTACAGCC	CAGGAGGGC	CGGCAGTGAC	1200
5 .	ACCTTGGTGA	GGTATTTTGC	AAATGAACCA	TTTGCTGACT	TCCACCGAGT	GGAAGGATTG	1260
	CAAGGAGTCT	ACATTGCTAC	TCTGATTAAT	GGTTCTATGA	ATGAGGAGAA	CATGAGATCG	1320
	GTCATCACCT	TTGACAAAGG	GGGAACCTGG	GAGTTTCTTC	AGGCTCCAGC	CTTCACGGGA	1380
10	TATGGAGAGA	AAATCAATTG	TGAGCTTTCC	CAGGGCTGTT	CCCTTCATCT	GGCTCAGCGC	1440
	CTCAGTCAGC	TCCTCAACCT	CCAGCTCCGG	AGAATGCCCA	TCCTGTCCAA	GGAGTCGGCT	1500
	CCAGGCCTCA	TCATCGCCAC	TGGCTCAGTG	GGAAAGAACT	TGGCTAGCAA	GACAAACGTG	1560
15	TACATCTCTA	GCAGTGCTGG	AGCCAGGTGG	CGAGAGGCAC	TTCCTGGACC	TCACTACTAC	1620
	ACATGGGGAG	ACCACGGCGG	AATCATCACG	GCCATTGCCC	AGGGCATGGA	AACCAACGAG	1680
	CTAAAATACA	GTACCAATGA	AGGGGAGACC	TGGAAAACAT	TCATCTTCTC	TGAGAAGCCA	1740
20	GTGTTTGTGT	ATGGCCTCCT	CACAGAACCT	GGGGAGAAGA	GCACTGTCTT	CACCATCTTT	1800
	GGCTCGAACA	AAGAGAATGT	CCACAGCTGG	CTGATCCTCC	AGGTCAATGC	CACGGATGCC	1860
	TTGGGAGTTC	CCTGCACAGA	GAATGACTAC	AAGCTGTGGT	CACCATCTGA	TGAGCGGGGG	1920
25	AATGAGTGTT	TGCTGGGACA	CAAGACTGTT	TTCAAACGGC	GGACCCCCCA	TGCCACATGC	1980
	TTCAATGGAG	AGGACTTTGA	CAGGCCGGTG	GTCGTGTCCA	ACTGCTCCTG	CACCCGGGAG	2040
	GACTATGAGT	GTGACTTCGG	TTTCAAGATG	AGTGAAGATT	TGTCATTAGA	GGTTTGTGTT	2100
30	CCAGATCCGG	AATTTTCTGG	AAAGTCATAC	TCCCCTCCTG	TGCCTTGCCC	TGTGGGTTCT	2160
00	ACTTACAGGA	GAACGAGAGG	CTACCGGAAG	ATTTCTGGGG	ACACTTGTAG	CGGAGGAGAT	2220
	GTTGAAGCGC	GACTGGAAGG	AGAGCTGGTC	CCCTGTCCCC	TGGCAGAAGA	GAACGAGTTC	2280
	ATTCTGTATG	CTGTGAGGAA	ATCCATCTAC	CGCTATGACC	TGGCCTCGGC	B AGCCACCGAG	2340
35	CAGTTGCCTC	TCACCGGGCT	ACGGGCAGCA	GTGGCCCTGG	ACTTTGACT	TGAGCACAAC	2400
	TGTTTGTATT	GGTCCGACCT	GGCCTTGGAC	GTCATCCAGO	GCCTCTGTT	r GAATGGAAGC	2460
	ACAGGGCAAG	AGGTGATCAT	CAATTCTGGC	CTGGAGACAG	TAGAAGCTT	r ggcttttgaa	2520
40	CCCCTCAGCC	AGCTGCTTT	CTGGGTAGAI	GCAGGCTTCA	AAAAGATTG	A GGTAGCTAAT	2580
	CCAGATGGCG	ACTTCCGACT	CACAATCGTO	AATTCCTCTC	TGCTTGATC	G TCCCAGGGCT	264
	CTGGTCCTCG	TGCCCCAAG	GGGGGTGAT	TTCTGGACA	ACTGGGGAG	A CCTGAAGCCT	270
45	GGGATTTATC	GGAGCAATA	GGATGGTTCI	GCTGCCTAT	CACCTGGTGT	C TGAGGATGTG	276
	AAGTGGCCCA	ATGGCATCT	TGTGGACGAC	CAGTGGATT	r actggacgg	A TGCCTACCTG	282
	GAGTGCATAG	AGCGGATCA	GTTCAGTGG	CAGCAGCGC	r ctgtcattc	T GGACAACCTC	288
50	CCGCACCCCT	ATGCCATTG	C TGTCTTTAA	AATGAAATC	r actgggatg	A CTGGTCACAG	294
	(M()) (()) (())		- CNNNTNCNC'	r GGGTCCCAG	A TGGAGATTC	T GGCAAACCAG	300

	CTCACGGGGC TCATGGACAT GAAGATTTTC TACAAGGGGA AGAACACTGG AAGCAATGCC	3060
	TGTGTGCCCA GGCCATGCAG CCTGCTGTGC CTGCCCAAGG CCAACAACAG TAGAAGCTGC	3120
5	AGGTGTCCAG AGGATGTGTC CAGCAGTGTG CTTCCATCAG GGGACCTGAT GTGTGACTGC	3120
	CCTCAGGGCT ATCAGCTCAA GAACAATACC TGTGTCAAAG AAGAGAACAC CTGTCTTCGC	3240
	AACCAGTATC GCTGCAGCAA CGGGAACTGT ATCAACAGCA TTTGGTGGTG TGACTTTGAC	3300
10	AACGACTGTG GAGACATGAG CGATGAGAGA AACTGCCCTA CCACCATCTG TGACCTGGAC	
	ACCCAGTTTC GTTGCCAGGA GTCTGGGACT TGTATCCCAC TGTCCTATAA ATGTGACCTT	3360
	GAGGATGACT GTGGAGACAA CAGTGATGAA AGTCATTGTG AAATGCACCA GTGCCGGAGT	3420
15	GACGAGTACA ACTGCAGTTC CGGCATGTGC ATCCGCTCCT CCTGGGTATG TGACGGGGAC	3480
	AACGACTGCA GGGACTGGTC TGATGAAGCC AACTGTACCG CCATCTATCA CACCTGTGAG	3540
	GCCTCCAACT TCCAGTGCCG AAACGGGCAC TGCATCCCCC AGCGGTGGGC GTGTGACGGG	3600
20	GATACGGACT GCCAGGATGG TTCCGATGAG GATCCAGTCA ACTGTGAGAA GAAGTGCAAT	3660
	GGATTCCGCT GCCCAAACGG CACTTGCATC CCATCCAGCA AACATTGTGA TGGTCTGCGT	3720
	GATTGCTCTG ATGGCTCCGA TGAACAGCAC TGCGAGCCCC TCTGTACGCA CTTCATGGAC	3780
<i>25</i>	TTTGTGTGTA AGAACCGCCA GCAGTGCCTG TTCCACTCCA TGGTCTGTGA CGGAATCATC	3840
23	CAGTGCCGCG ACGGGTCCGA TGAGGATGCG GCGTTTGCAG GATGCTCCCA AGATCCTGAG	3900
	TTCCACAAGG TATGTGATGA GTTCGGTTTC CAGTGTCAGA ATGGAGTGTG CATCAGTTTG	3960
	ATTTGGAAGT GCGACGGGAT GGATGATTGC GGCGATTATT CTGATGAAGC CAACTGCGAA	4020
30	AACCCCACAG AAGCCCCAAA CTGCTCCCGC TACTTCCAGT TTCGGTGTGA GAATGGCCAC	4080
	TGCATCCCCA ACAGATGGAA ATGTGACAGG GAGAACGACT GTGGGGGACTG GTCTGATGAG	4140
	AAGGATTGTG GAGATTCACA TATTCTTCCC TTCTCGACTC CTGGGCCCTC CACGTGTCTG	4200
35	\cdot	4260
	CCCAATTACT ACCGCTGCAG CAGTGGGACC TGCGTGATGG ACACCTGGGT GTGCGACGGG TACCGAGATT GTGCAGATGG GTGTGAGGAG GAAGGGTGGG	4320
	TACCGAGATT GTGCAGATGG CTCTGACGAG GAAGCCTGCC CCTTGCTTGC AAACGTCACT	4380
40	GCTGCCTCCA CTCCCACCCA ACTTGGGCGA TGTGACCGAT TTGAGTTCGA ATGCCACCAA	4440
	CCGAAGACGT GTATTCCCAA CTGGAAGCGC TGTGACGGCC ACCAAGATTG CCAGGATGGC	4500
	CGGGACGAGG CCAATTGCCC CACACAGGC ACCTTGACTT GCATGAGCAG GGAGTTCCAG	4560
4 5	TGCGAGGACG GGGAGGCCTG CATTGTGCTC TCGGAGCGCT GCGACGGCTT CCTGGACTGC	4620
	TCGGACGAGA GCGATGAAAA GGCCTGCAGT GATGAGTTGA CTGTGTACAA AGTACAGAAT	4680
	CTTCAGTGGA CAGCTGACTT CTCTGGGGAT GTGACTTTGA CCTGGATGAG GCCCAAAAAA	4740
50	ATGCCCTCTG CATCTTGTGT ATATAATGTC TACTACAGGG TGGTTGGAGA GAGCATATGG	4800
	AAGACTCTGG AGACCCACAG CAATAAGACA AACACTGTAT TAAAAGTCTT GAAACCAGAT	4860
	ACCACGTATC AGGTTAAAGT ACAGGTTCAG TGTCTCAGCA AGGCACAACAA CACCAATGAC	4920

	TTTGTGACCC	TGAGGACCCC	AGAGGGATTG	CCAGATGCCC	CTCGAAATCT	CCAGCTGTCA	4980
	CTCCCCAGGG	AAGCAGAAGG	TGTGATTGTA	GGCCACTGGG	CTCCTCCCAT	CCACACCCAT	5040
5	GGCCTCLTCC	GTGAGTACAT	TGTAGAATAC	AGCAGGAGTG	GTTCCAAGAT	GTGGGCCTCC	5100
	CAGAGGGCTG	CTAGTAACTT	TACAGAAATC	AAGAACTTAT	TGGTCAACAC	TCTATACACC	5160
	GTCAGAGTGG	CTGCGGTGAC	TAGTCGTGGA	ATAGGAAACT	GGAGCGATTC	TAAATCCATT	5220
10	ACCACCATAA	AAGGAAAAGT	GATCCCACCA	CCAGATATCC	ACATTGACAG	CTATGGTGAA	5280
	AATTATCTAA	GCTTCACCCT	GACCATGGAG	AGTGATATCA	AGGTGAATGG	CTATGTGGTG	5340
	AACCTTTTCT	GGGCATTTGA	CACCCACAAG	CAAGAGAGGA	GAACTTTGAA	CTTCCGAGGA	5400
15	AGCATATTGT	CACACAAAGT	TGGCAATCTG	ACAGCTCATA	CATCCTATGA	GATTTCTGCC	5460
	TGGGCCAAGA	CTGACTTGGG	GGATAGCCCT	CTGGCATTTG	AGCATGTTAT	GACCAGAGGG	5520
	GTTCGCCCAC	CTGCACCTAG	CCTCAAGGCC	AAAGCCATCA	ACCAGACTGC	AGTGGAATGT	5580
20	ACCTGGACCG	GCCCCGGAA	TGTGGTTTAT	GGTATTTTCT	ATGCCACGTC	CTTTCTTGAC	5640
	CTCTATCGCA	ACCCGAAGAG	CTTGACTACT	TCACTCCACA	ACAAGACGGT	CATTGTCAGT	5700
	AAGGATGAGC	AGTATTTGTT	TCTGGTCCGT	GTAGTGGTAC	CCTACCAGGG	GCCATCCTCT	5760
25	GACTACGTTG	TAGTGAAGAT	GATCCCGGAC	AGCAGGCTTC	CACCCCGTCA	CCTGCATGTG	5820
	GTTCATACGG	GCAAAACCTC	CGTGGTCATC	AAGTGGGAAT	CACCGTATGA	CTCTCCTGAC	5880
	CAGGACTTGT	TGTATGCAAT	TGCAGTCAAA	GATCTCATAA	GAAAGACTGA	CAGGAGCTAC	5940
30	AAAGTAAAAT	CCCGTAACAG	CACTGTGGAA	TACACCCTTA	ACAAGTTGGA	GCCTGGCGGG	6000
30	AAATACCACA	TCATTGTCCA	ACTGGGGAAC	ATGAGCAAAG	ATTCCAGCAT	AAAAATTACC	6060
	ACAGTTTCAT	TATCAGCACC	TGATGCCTTA	AAAATCATAA	CAGAAAATGA	TCATGTTCTT	6120
	CTGTTTTGGA	AAAGCCTGGC	TTTAAAGGAA	AAGCATTTTA	ATGAAAGCAG	GGGCTATGAG	6180
35	ATACACATGT	TTGATAGTGC	CATGAATATC	ACAGCTTACC	TTGGGAATAC	TACTGACAAT	6240
	TTCTTTAAAA	TTTCCAACCI	GAAGATGGGT	CATAATTACA	CGTTCACCGT	CCAAGCAAGA	6300
	TGCCTTTTTG	GCAACCAGAT	CTGTGGGGAG	CCTGCCATCC	TGCTGTACGA	TGAGCTGGGG	6360
40	TCTGGTGCAG	ATGCATCTGC	AACGCAGGCT	GCCAGATCTA	CGGATGTTG	TGCTGTGGTG	6420
	GTGCCCATCT	TATTCCTGAT	ACTGCTGAGO	CTGGGGGTGG	GGTTTGCCAT	CCTGTACACG	6480
	AAGCACCGGA	GGCTGCAGAG	CAGCTTCACC	GCCTTCGCCA	ACAGCCACT	A CAGCTCCAGG	6540
45	CTGGGGTCCG	CAATCTTCT	CTCTGGGGA	GACCTGGGGG	AAGATGATG	A AGATGCCCCT	6600
	ATGATAACTG	GATTTTCAG	A TGACGTCCCC	ATGGTGATAG	CC		6643

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2214 amino acids
(B) TYPE: amino acid

55

(C) STRANDEDNESS:
(D) TOPOLOGY: linear

5	(ii)	MOI	ECUI	LE TY	PE:	Prot	ein								•	
	(xi)	SEÇ	UENC	E DE	SCRI	PTIC	N: S	SEQ I	D NC): 6:						
10		Ala										J Leu	Pro	Phe	Leu 15	Phe
•	Thr	Leu	Val	Ala 20	Leu	Leu	Pro	Pro	Gly 25	Ala	Leu	. Cys	Glu	. Val	Trp	Thi
15	Gln	Arg	Leu 35	His	Gly	Gly	Ser	Ala 40	Pro	Leu	Pro	Gln	Asp 45	Arg	Gly	Phe
	Leu	Val 50	Val	Gln	Gly	Asp	Pro 55	Arg	Glu	Leu	Arg	Leu 60	Trp	Ala	Arg	Gly
20		Ala				70					75					80
		Ser			0,5					90					95	
25		Leu		100					105					110		
		Ser						120					125			•
30		Pro 130					133					140				
		Lys				130					155					160
35		Ala			105					170					175	
		Ile		100					182					190		
40		Cys	1))					200					205			
		Leu 210					215					220				
45		His				250					235					240
•		Ile			243					250					255	
50		Tyr		200					265					270		
	Gly	Tyr	Ser 275	Thr	Val	Phe	Arg	Ser 280	Thr	Asp	Phe	Phe	Gln 285	Ser	Arg	Glu

90

	Asn	Gln 290	Glu	Val	Ile	Leu	Glu 295	Glu	Val	A.rg	Asp	300 5µ5	Cln	Leu	Arg	Asp
5	Lys 305	Tyr	Met	Phe	Ala	Thr 310	Lys	Val	Val	His	Leu 315	Leu	Gly	Ser	Glu	Gln 320
	Gln	Ser	Ser	Val	Gln 325	Leu	Trp	Val	Ser	Phe 330	Gly	Arg	Lys	Pro	Met 335	Arg
10	Ala	Ala	Gln	Phe 340	Val	Thr	Arg	His	Pro 345	Ile	Asn	Glu	Tyr	Tyr 350	Ile	Ala
	Asp	Ala	Ser 355	Glu	Asp	Gln	Val	Phe 360	Val	Сув	Val	Ser	His 365	Ser	Asn	Asn
15	Arg	Thr 370	Asn	Leu	Tyr	Ile	Ser 375	Glu	Ala	Glu	Gly	Leu 380	Lys	Phe	Ser	Leu
	Ser 385	Leu	Glu	Asn	Val	Leu 390	Tyr	Tyr	Ser	Pro	Gly 395	Gly	Ala	Gly	Ser	Asp 400
20	Thr	Leu	Val	Arg	Tyr 405	Phe	Ala	Asn	Glu	Pro 410	Phe	Ala	Asp	Phe	His 415	Arg
	Val	Glu	Gly	Leu 420	Gln	Gly	Val	Tyr	Ile 425	Ala	Thr	Leu	Ile	Asn 430	Gly	Ser
25	Met	Asn	Glu 435	Glu	Asn	Met	Arg	Ser 440	Val	Ile	Thr	Phe	Asp 445	Lys	Gly	Gly
	Thr	Trp 450	Glu	Phe	Leu	Gln	Ala 455	Pro	Ala	Phe	Thr	Gly 460	Tyr	Gly	Glu	Lys
3 <i>0</i>	Ile 465	Asn	Cys	Glu	Leu	Ser 470	Gln	Gly	Cys	Ser	Leu 475	His	Leu	Ala	Gln	Arg 480
	Leu	Ser	Gln	Leu	Leu 485	Asn	Leu	Gln	Leu	Arg 490	Arg	Met	Pro	Ile	Leu 495	Ser
35	Lys	Glu	Ser	Ala 500	Pro	Gly	Leu	Ile	Ile 505	Ala	Thr	Gly	Ser	Val 510	Gly	Lys
-	Asn	Leu	Ala 515	Ser	Lys	Thr	Asn	Val 520	Tyr	Ile	Ser	ser	Ser 525	Ala	Gly	Ala
40	Arg	Trp 530	Arg	Glu	Ala	Leu	Pro 535	Gly	Pro	His	Tyr	Tyr 540	Thr	Trp	Gly	Asp
40	His 5 4 5	Gly	Gly	Ile	Ile	Thr 550	Ala	Ile	Ala	Gln	Gly 555	Met	Glu	Thr	Asn	Glu 560
	Leu	Lys	Tyr	Ser	Thr 565	Asn	Glu	Gly	Glu	Thr 570	Trp	Lys	Thr	Phe	Ile 575	Phe
45	Ser	Glu	Lys	Pro 580	Val	Phe	Val	Tyr	Gly 585	Leu	Leu	Thr	Glu	Pro 590	Gly	Glu
	Lys	Ser	Thr 595	Val	Phe	Thr	Ile	Phe 600	Gly	Ser	Asn	Lys	Glu 605	Asn	Val	His
50	Ser	Trp 610	Leu	Ile	Leu	Gln	Val 615		Ala	Thr	Asp	Ala 620	Leu	Gly	Val	Pro
	Cys	Thr	Glu	Asn	Asp	Tyr	Lys	Leu	Trp	Ser	Pro	Ser	Asp	Glu	Arg	Gly

	625					630					635					640
5	Asn	Glu	Cys	Leu	Leu 645	Gly	His	Lys	Thr	Val 650	Phe	Lys	Arg	_	Thr 655	Pro
	His	Ala	Thr	Cys 660	Phe	Asn	Gly	Glu	Asp 665	Phe	Asp	Arg	Pro	Val 670	Val	Val
10	Ser	Asn	Суз 675	Ser	Сув	Thr	Arg	Glu 680	Asp	Tyr	Glu	Cys	Asp 685	Phe	Gly	Phe
	Lys	Met 690	Ser	Glu	Asp	Leu	Ser 695	Leu	Glu	Val	Сув	Val 700	Pro	Asp	Pro	Glu
15	Phe 705	Ser	Gly	Lys	Ser	Tyr 710	Ser	Pro	Pro	Val	Pro 715	Суз	Pro	Val	Gly	Ser 720
15	Thr	Tyr	Arg	Arg	Thr 725	Arg	Gly	Tyr	Arg	Lys 730	Ile	Ser	Gly	Asp	Thr 735	Cys
	Ser	Gly	Gly	Asp 740	Val	Glu	Ala	Arg	Leu 745	Glu	Gly	Glu	Leu	Val 750	Pro	Cys
20	Pro	Leu	Ala 755	Glu	Glu	Asn	Glu	Phe 760	Ile	Leu	Tyr	Ala	Val 765	Arg	Lys	Ser
	lle	Tyr 770	Arg	Tyr	Asp	Leu	Ala 775	Ser	Gly	Ala	Thr	Glu 780	Gln	Leu	Pro	Leu
25	Thr 785	Gly	Leu	Arg	Ala	Ala 790	Val	Ala	Leu	Asp	Phe 795	Asp	Tyr	Glu	His	Asn 800
	Cys	Leu	Tyr	Trp	Ser 805	Asp	Leu	Ala	Leu	Asp 810	Val	Ile	Gln	Arg	Leu 815	Cys
30	Leu	Asn	Gly	Ser 820	Thr	Gly	Gln	Glu	Val 825	Ile	Ile	Asn	Ser	Gly 830	Leu	Glu
	Thr	Val	Glu 835	Ala	Leu	Ala	Phe	Glu 840	Pro	Leu	Ser	Gln	Leu 845	Leu	Tyr	Trp
35	Val	Asp 850	Ala	Gly	Phe	Lys	Lys 855	Ile	Glu	Val	Ala	Asn 860	Pro	Asp	Gly	Asp
	Phe 865	Arg	Leu	Thr		Val 870	Asn	Ser	Ser	Val	Leu 875	Asp	Arg	Pro	Arg	Ala 880
40	Leu	Val	Leu	Val	Pro 885	Gln	Glu	Gly	Val	Met 890	Phe	Trp	Thr	Asp	Trp 895	Gly
	Asp	Leu	Lys	Pro 900	Gly	Ile	Tyr	Arg	Ser 905	Asn	Met	Asp	Gly	Ser 910	Ala	Ala
45	Tyr	His	Leu 915	Val	Ser	Glu	Asp	Val 920	Lys	Trp	Pro	Asn	Gly 925	Ile	Ser	Val
	Asp	Asp 930	Gln	Trp	Ile	Tyr	Trp 935	Thr	Asp	Ala	Tyr	Leu 940	Glu	Сув	Ile	Glu
50	Arg 945	Ile	Thr	Phe	Ser	Gly 950	Gln	Gln	Arg	Ser	Val 955	Ile	Leu	Asp	Asn	Leu 960
	Pro	His	Pro	Tyr	Ala 965	Ile	Ala	Val	Phe	Lys 970	Asn	Glu	Ile	Tyr	Trp 975	Asp

	_		_		_	_			_							
	Asp	Trp	Ser	Gln 980	Leu	Ser	Ile	Phe	A rg 985	Ala	Sex	Lуя	Tyr	Ser 990	Gly	Ser
5	Gln	Met	Glu 995	Ile	Leu	Ala	Asn	Gln 1000		Thr	Gly	Leu	Met 1005		Met	Lys
	Ile	Phe 1010		Lys	Gly	Lys	Asn 1015		Gly	Ser	Asn	Ala 1020		Val	Pro	Arg
10	Pro 1025		Ser	Leu	Leu	Суз 1030	Leu)	Pro	Lys	Ala	Asn 1035		Ser	Arg	Ser	Cys 1040
	Arg	Cys	Pro	Glu	Asp 1045		Ser	Ser	Ser	Val 1050		Pro	Ser	Gly	Asp 1055	
15	Met	Cys	Asp	Cys 1060		Gln	Gly	Tyr	Gln 1065		Lys	Asn	Asn	Thr 1070		Val
	Lys	Glu	Glu 1075		Thr	Cys	Leu	Arg 1080		Gln	Tyr	Arg	Cys 1085		Asn	Gly
20	Asn	Cys 1090		Asn	Ser	Ile	Trp 1095	_	Сув	Asp	Phe	Asp 1100		Asp	Cys	Gly
	Asp 1105		Ser	Asp	Glu	Arg 1110		Суз	Pro	Thr	Thr 1115		Cys	Asp	Leu	Asp 1120
25	Thr	Gln	Phe	Arg	Cys 1125		Glu	Ser	Gly	Thr 1130		Ile	Pro	Leu	Ser 1135	
	Lys	Cys	Asp	Leu 1140		Asp	Asp	Cys	Gly 1145		Asn	Ser	Asp	Glu 1150		His
30	Cys	Glu	Met 1155		Gln	Cys	Arg	Ser 1160		Glu	Tyr	Asn	Cys 1165		Ser	Gly
	Met	Cys 1170		Arg	Ser	Ser	Trp 1175		Cys	Asp	Gly	Asp 1180		Asp	Cys	Arg
35	Asp 1185		Ser	Asp	Glu	Ala 1190		Cys	Thr	Ala	Ile 1195	_	His	Thr	Cys	Glu 1200
	Ala	Ser	Asn	Phe	Gln 1205		Arg	Asn	Gly	His 1210		Ile	Pro	Gln	Arg 1215	
40	Ala	Cys	Asp	Gly 1220		Thr	Asp	Cys	Gln 1225		Gly	Ser	Asp	Glu 1230		Pro
· -	Val	Asn	Cys 1235		Lys	Lys	Cys	Asn 1240		Phe	Arg	Cys	Pro 1245		Gly	Thr
	Cys	Ile 1250		Ser	Ser	Lys	His 1255		Asp	Gly	Leu	Arg 1260	_	Сув	Ser	Asp
45	Gly 1265		Asp	Glu	Gln	His 1270		Glu	Pro	Leu	Cys 1275		His	Phe	Met	Asp 1280
	Phe	Val	Сув	Lys	Asn 1285		Gln	Gln	Суз	Leu 1290		His	Ser	Met	Val 1295	
50	Asp	Gly	Ile	Ile 1300		Cys	Arg	Asp	Gly 1305		Asp	Glu	Asp	Ala 1310	Ala)	Phe

	Ala	Gly	Cys 1319	Ser 5	Gln	Asp	Pro	Glu 1320	Phe)	His	Lys	Ľs"	Сув 1329		Clu	Ph≘
5	Gly	Phe 1330	Gln O	Cys	Gln	Asn	Gly 1335	Val	Cys	Ile	Ser	Leu 1340		Trp	Lys	Cys
	Asp 1345	Gly 5	Met	Asp	Asp	Cys 1350	Gly)	Asp	Tyr	Ser	Asp 1355		Ala	Asn	Суз	Glu 1360
10	Asn	Pro	Thr	Glu	Ala 1369	Pro	Asn	Cys	Ser	Arg 1370		Phe	Gln	Phe	Arg 1379	
	Glu	Asn	Gly	His 1380	Cys)	Ile	Pro	Asn	Arg 1389		Lys	Суз	Asp	Arg 1390		Asn
15	Asp	Суз	Gly 1395	Asp	Trp	Ser	Asp	Glu 1400	Lys)	Asp	Cys	Gly	Asp 1405		His	Ile
	Leu	Pro 1410	Phe)	Ser	Thr	Pro	Gly 1415	Pro	Ser	Thr	Суз	Leu 1420		Asn	Tyr	Tyr
20	Arg 1425	Cys	Ser	Ser	Gly	Thr 1430	Cys)	Val	Met	Asp	Thr 1439		Val	Cys	Asp	Gly 1440
	Tyr	Arg	Asp	Cys	Ala 1449	Asp	Gly	Ser	Asp	Glu 1450		Ala	Cys	Pro	Leu 1455	
25	Ala	Asn	Val	Thr 1460	Ala	Ala	Ser	Thr	Pro 1465		Gln	Leu		Arg 1470		Asp
	Arg	Phe	Glu 1475	Phe	Glu	Суз	His	Gln 1480		Lys	Thr	Суз	Ile 1485		Asn	Trp
30	Lys	Arg 1490	Cys)	Asp	Gly	His	Gln 1495		Cys	Gln	Asp	Gly 1500		Asp	Glu	Ala
	Asn 1505	Cys	Pro	Thr	His	Ser 1510	Thr	Leu	Thr	Суз	Met 1515		Arg	Glu	Phe	Gln 1520
35	Cys	Glu	Asp	Gly	Glu 1525		Cys	Ile	Val	Leu 1530		Glu	Arg	Сув	Asp 1535	_
	Phe	Leu	Asp	Cys 1540		Asp	Glu	Ser	Asp 1545		Lys	Ala	Cys	Ser 1550	_	Glu
40	Leu	Thr	Val 1555	Tyr	Lys	Val	Gln	Asn 1560	Leu)	Gln	Trp	Thr	Ala 1565		Phe	Ser
	Gly	Asp 1570	Val	Thr	Leu	Thr	Trp 1575	Met 5	Arg	Pro	Lys	Lys 1580		Pro	Ser	Ala
	Ser 1585	Cys	Val	Tyr	Asn	Val 1590	Tyr)	Tyr	Arg	Val	Val 1599		Glu	Ser	Ile	Trp 1600
45	Lys	Thr	Leu	Glu	Thr 1605		Ser	Asn	Lys	Thr 1610		Thr	Val	Leu	Lys 1615	
	Leu	Lys	Pro	Asp 1620	Thr	Thr	Tyr	Gln	Val 1625		Val	Gln	Val	Gln 1630		Leu
50	Ser	Lys	Ala 1635		Asn	Thr	Asn	Asp 1640		Val	Thr	Leu	Arg 1645		Pro	Glu
	Gly	Leu	Pro	Asp	Ala	Pro	Arg	Asn	Leu	Gln	Leu	Ser	Leu	Pro	Arg	Glu

		1650	i				1655	;				3.650)			
;	Ala 1665		Gly	Val	Ile	Val 1670	Gly	His	Trp	Ala	Pro 1675		Ile	His	Thr	His 1680
	Gly	Leu	Ile	Arg	Glu 1685		Ile	Val	Glu	Tyr 1690		Arg	Ser	Gly	Ser 1695	•
10	Met	Trp	Ala	Ser 1700		Arg	Ala	Ala	Ser 1705		Phe	Thr	Glu	Ile 1710		Asn
•	Leu	Leu	Val 1715		Thr	Leu	Tyr	Thr 1720		Arg	Val	Ala	Ala 1725		Thr	Ser
	Arg	Gly 1730		Gly	Asn	Trp	Ser 1735		Ser	Lys	Ser	Ile 1740		Thr	Ile	Lys
15	Gly 1745	-	Val	Ile	Pro	Pro 1750	Pro	qaA	Ile	His	Ile 1755		Ser	Tyr	Gly	Glu 1760
	Asn	Tyr	Leu	Ser	Phe 1765		Leu	Thr	Met	Glu 1770		Asp	Ile	Lys	Val 1775	
20	Gly	Tyr	Val	Val 1780		Leu	Phe	Trp	Ala 1785		Asp	Thr	His	Lys 1790		Glu
	Arg	Arg	Thr 1795		Asn	Phe	Arg	Gly 1800		Ile	Leu	Ser	His 1805		Val	Gly
25	Asn	Leu 1810		Ala	His	Thr	Ser 1815		Glu	Ile	Ser	Ala 1820		Ala	Lys	Thr
	Asp 1825		Gly	Asp	Ser	Pro 1830		Ala	Phe	Glu	His 1835		Met	Thr	Arg	Gly 1840
30	Val	Arg	Pro	Pro	Ala 1845		Ser	Leu	Lys	Ala 1850		Ala	Ile	Asn	Gln 1855	
-	Ala	Val	Glu	Cys 1860		Trp	Thr	Gly	Pro 1869		Asn	Val	Val	Tyr 1870		Ile
35	Phe	Tyr	Ala 1879		Ser	Phe	Leu	Asp 1880		Tyr	Arg	Asn	Pro 1885		Ser	Leu
	Thr	Thr 1890		Leu	His	Asn	Lys 1899		Val	Ile	Val	Ser 190		Asp	Glu	Gln
40	Tyr 1905		Phe	Leu	Val	Arg 1910		Val	Val	Pro	Tyr 191		Gly	Pro	Ser	Ser 1920
	Asp	Tyr	Val	Val	Val 1929		Met	Ile	Pro	Asp 193		Arg	Leu	Pro	Pro 1935	
45	His	Leu	His	Val 1940		His	Thr	Gly	Lys 194		Ser	Val	Val	Ile 195		Trp
	Glu	Ser	Pro 195		Asp	Ser	Pro	Asp 1960		Asp	Leu	Leu	Tyr 196		Ile	Ala
50	Val	Lys 1970		Leu	Ile	Arg	Lys 197		Asp	Arg	Ser	Tyr 198		Val	Lys	Ser
	Arg 1985		Ser	Thr	Val	Glu 199		Thr	Leu	Asn	Lys 199		Glu	Pro	Gly	Gly 2000

		Lys	Tyr	His	Ile	Ile 200	Val	Gln	Leu	Сlу	Asn 2010	Met O	Ser	Lys	Asp	Ser 201	
5		Ile	Lys	Ile	Thr 202	Thr 0	Val	Ser	Leu	Ser 2025	Ala	Pro	Asp	Ala	Leu 203		Ile
		Ile	Thr	Glu 203	Asn 5	qaA	His	Val	Leu 2040	Leu	Phe	Trp	Lys	Ser 204		Ala	Leu
10		Lys	Glu 205	Lys 0	His	Phe	Asn	Glu 205	Ser	Arg	Gly	Tyr	Glu 2060		His	Met	Phe
		Asp 2065	Ser 5	Ala	Met	Asn	Ile 2070	Thr	Ala	Tyr	Leu	Gly 2075	Asn	Thr	Thr	Asp	Asn 2080
15		Phe	Phe	Lys	Ile	Ser 208	Asn	Leu	Lys	Met	Gly 2090		Asn	Tyr	Thr	Phe 2095	
		Val	Gln	Ala	Arg 2100	Cys	Leu	Phe	Gly	Asn 2105		Ile	Cys	Gly	Glu 2110		Ala
20		Ile	Leu	Leu 2119	Tyr	Asp	Glu	Leu	Gly 2120	Ser	Gly	Ala	Asp	Ala 2125		Ala	Thr
		Gln	Ala 2130	Ala	Arg	Ser	Thr	Asp 2135		Ala	Ala	Val	Val 2140		Pro	Ile	Leu
25		Phe 2145	Leu	Ile	Leu	Leu	Ser 2150	Leu)	Gly	Val	Gly	Phe 2155		Ile	Leu	Tyr	Thr 2160
		Lys	His	Arg	Arg	Leu 2165	Gln	Ser	Ser	Phe	Thr 2170		Phe	Ala	Asn	Ser 2175	
30		Tyr	Ser	Ser	Arg 2180	Leu)	Gly	Ser	Ala	Ile 2185		Ser	Ser	Gly	Asp 2190		Leu
		Gly	Glu	Asp 2195	Asp	Glu	Asp	Ala	Pro 2200		Ile	Thr	Gly	Phe 2205		Asp	Asp
35		Val	Pro 2210	Met)	Val	Ile	Ala										
	(2)	INFOR	ITAM	ON F	OR S	SEQ I	D NC): 7:									
4 0		(i)	(A) (B)	JENCE LEN TYF STR TOF	IGTH: PE: n	684 nucle	3 ba ic a SS: d	se p cid loubl	airs	i							
		(ii)	MOLE	CULE	TYF	E: c	DNA	to m	RNA								
45		(ix)	(A) (B)	TURE: NAM LOC IDE	E/KE	N:81	16	4		s							
50		(ix)	(A) (B)	URE: NAM LOC IDE	E/KE	N:16	56	722		s							
55																	

(xi) SEQUENCE DESCRIPTION: SEQ ID MO: 1:

5	TCG	CGCT	GCA (CATT	CTCT	CC TO	GCG	GCGG	C GC	CACC'	CCG IGCA	GCC(GTA(CAGC(GCGT	GGC TCG	TCTCG	TTGGCC AACATG Met	23 83
	GCG Ala	ACA Thr	CGG Arg	AGC Ser	AGC Ser	AGG Arg	AGG Arg	GAG Glu	TCG Ser 10	CGA Arg	CTC Leu	CCG Pro	TTC Phe	CTA Leu 15	TTC Phe	ACC Thr	131
10	Leu	Val	Ala 20	Leu	Leu	Pro	Pro	Gly 25	Ala	Leu	Cys	Glu	Val 30	TGG Trp	ACG Thr	Gln	179
	Arg	Leu 35	His	Gly	Gly	Ser	Ala 40	Pro	Leu	Pro	Gln	Asp 45	Arg	Gly	TTC Phe	Leu	227
15	Val 50	Val	Gln	Gly	Asp	Pro 55	Arg	Glu	Leu	Arg	Leu 60	Trp	Ala	Arg	G GG Gly	Asp 65	275
	Ala	Arg	Gly	Ala	Ser 70	Arg	Ala	Asp	Glu	Lys 75	Pro	Leu	Arg	Arg	AAA Lys 80	Arg	323
20	Ser	Ala	Ala	Leu 85	Gln	Pro	Glu	Pro	Ile 90	Lys	Val	Tyr	Gly	Gln 95	GTT Val	Ser	371
	Leu	Asn	Asp 100	Ser	His	Asn	Gln	Met 105	Val	Val	His	Trp	Ala 110	Gly	GAG Glu	Lys	419
25	Ser	Asn 115	Val	Ile	Val	Ala	Leu 120	Ala	Arg	Asp	Ser	Leu 125	Ala	Leu	GCG Ala	Arg	467
	Pro 130	Lys	Ser	Ser	Asp	Val 135	Tyr	Val	Ser	Tyr	Asp 140	Tyr	Gly	Lys	TCA Ser	Phe 145	515
30	AAG Lys	Lys	ATT Ile	TCA Ser	GAC Asp 150	AAG Lys	TTA Leu	AAC Asn	TTT Phe	GGC Gly 155	TTG Leu	GGA Gly	AAT Asn	AGG Arg	AGT Ser 160	GAA Glu	563
	Ala	Val	Ile	Ala 165	Gln	Phe	Tyr	His	Ser 170	Pro	Ala	Asp	Asn	Lys 175	CGG Arg	Tyr	611
35	Ile	Phe	Ala 180	Asp	Ala	Tyr	Ala	Gln 185	Tyr	Leu	Trp	Ile	Thr 190	Phe	GAC Asp	Phe	659
	Cys	Asn 195	Thr	Leu	Gln	Gly	Phe 200	Ser	Ile	Pro	Phe	Arg 205	Ala	Ala	GAT Asp	Leu	707
40	Leu 210	Leu	His	Ser	Lys	Ala 215	Ser	Asn	Leu	Leu	Leu 220	Gly	Phe	Asp	AGG Arg	Ser 225	755
	His	Pro	Asn	Lys	Gln 230	Leu	Trp	Lys	Ser	Asp 235	Asp	Phe	Gly	Gln	ACC Thr 240	Trp	803
	Ile	Met	Ile	Gln 245	Glu	His	Val	Lys	Ser 250	Phe	Ser	Trp	Gly	Ile 255	GAT Asp	Pro	851
45	Tyr	Asp	Lys 260	Pro	Asn	Thr	Ile	Tyr 265	Ile	Glu	Arg	His	Glu 270	Pro	TCT Ser	Gly	899
	Tyr	Ser 275	Thr	Val	Phe	Arg	Ser 280	Thr	Asp	Phe	Phe	Gln 285	Ser	Arg	GAA Glu	Asn	947
. 50	CAG Gln 290	GAA Glu	GTG Val	ATC Ile	CTT Leu	GAG Glu 295	GAA Glu	GTG Val	AGA Arg	GAT Asp	TTT Phe 300	CAG Gln	CTT Leu	CGG Arg	GAC Asp	AAG Lys 305	995

	TAC Tyr	ATG Met	TTT Phe	GCT Ala	ACA Thr 310	AAG Lys	GTG Val	GTG Val	CAT His	CTC Leu 315	TTG Leu	GGC GGC	ACT Ser	GAA Glu	CAC Gln 320	CAG Gln	1943
5	TCT Ser	TCT Ser	GTC Val	CAG Gln 325	CTC Leu	TGG Trp	G.LC A'1	TCC Ser	TTT Phe 330	GGC	CGG Arg	AAG Lys	CCC Pro	ATG Met 335	AGA	GCA Ala	1091
	Ala	Gln	Phe 340	GTC Val	Thr	Arg	His	CCT Pro 345	ATT Ile	Asn	Glu	Tyr	Tyr 350	ATC Ile	Ala	Asp	1139
10								GTG Val									1187
	Thr 370	Asn	Leu	Tyr	Ile	Ser 375	Glu	GCA Ala	Glu	Gly	Leu 380	Lys	Phe	Ser	Leu	Ser 385	1235
15								AGC Ser									1283
								GAA Glu									1331
20					_	_		ATT Ile 425	_	_		_		_			1379
	Asn	Glu 435	Glu	Asn	Met	Arg	Ser 440	GTC Val	Ile	Thr	Phe	Asp 445	Lya	Gly	Gly	Thr	1427
25	Trp 450	Glu	Phé	Leu	Gln	Ala 455	Pro	GCC Ala	Phe	Thr	Gly 460	Tyr	Gly	Glu	Lys	Ile 465	1475
								TGT Cys									1523
30								CTC Leu									1571
								ATC Ile 505									1619
35	Leu	Ala 515	Ser	Lys	Thr	Asn	Val 520	Tyr	Ile	Ser	Ser	Ser 525	Ala	Gly	Ala		1667
	Trp 530	Arg	Glu	Ala	Leu	Pro 535	Gly	CCT Pro	His	Tyr	Tyr 540	Thr	Trp	Gly	Asp	His 545	1715
4 0	Gly	Gly	Ile	Ile	Thr 550	Ala	Ile	GCC Ala	Gln	Gly 5 55	Met	Glu	Thr	Asn	Glu 560	Leu	1763
	Lys	Tyr	Ser	Thr 565	Asn	Glu	Gly	GAG Glu	Thr 570	Trp	Lys	Thr	Phe	Ile 575	Phe	Ser	1811
45	Glu	Lys	Pro 580	Val	Phe	Val	Tyr	GGC Gly 585	Leu	Leu	Thr	Glu	Pro 590	Gly	Glu	Lys	1859
45	Ser	Thr 595	Val	Phe	Thr	Ile	Phe 600		Ser	Asn	Lys	Glu 605	Asn	Val	His	Ser	1907
	Trp 610	Leu	Ile	Leu	Gln	Val 615	Asn	GCC Ala	Thr	Asp	Ala 620	Leu	Gly	Val	Pro	Cys 625	1955
50	Thr	Glu	Asn	Asp	Tyr 630	Lys	Leu	TGG Trp	Ser	Pro 635	Ser	Asp	Glu	Arg	Gly 640	Asn	2003
	GAG	TGT	TTG	CTG	GGA	CAC	AAG	ACT	GTT	TTC	AAA	CGG	CGG	ACC	CCC	CAT	2051

				043	•				650	1				655		His	
5	GCC Ala	ACA Thr	TGC Cys 660	File	TAA :	GGA Gly	GAG Glu	Asp	TTT Phe	GAC	AGG Arg	CCG Pro	Val	CTC	CMC	TCC Ser	2099
	AAC Asn	TGC Cys 675	TCC Ser	TGC	ACC Thr	CGG Arg	GIu	Asp	TAT	GAG Glu	TGT Cys	Asp	670 TTC Phe	GGT Gly	TTC Phe	AAG Lys	2147
10	ATG Met 690	AGT Ser	GAA	GAT Asp	TTG Leu	Ser	680 TTA Leu	GAG Glu	GTT Val	TGT Cys	Val	685 CCA Pro	GAT Asp	CCG Pro	GAA Glu	TTT	2195
	TCT	GGA	AAG Lys	TCA Ser	TAC Tyr 710	695 TCC Ser	CCT Pro	CCT Pro	GTG Val	Pro	700 TGC Cys	CCT Pro	GTG Val	GGT Gly	Ser	Thr	2243
15	TAC Tyr	AGG Arg	AGA Arg	ACG Thr 725	AGA	GGC Gly	TAC Tyr	CGG Arg	Lys	715 ATT Ile	TCT Ser	GGG Gly	GAC Asp	Thr	720 TGT Cys	NGC	2291
,5	GGA Gly	GGA Gly	GAT Asp 740	GTT	GAA Glu	GCG Ala	CGA Arg	CTG Leu 745	730 GAA Glu	GGA Gly	GAG Glu	CTG Leu	Val	735 CCC Pro	TGT Cys	CCC Pro	2339
	C T G Leu	GCA Ala 755	GAA	GAG Glu	AAC Asn	GAG Glu	TTC Phe 760	ATT	CTG Leu	TAT Tyr	GCT Ala	Val	750 AGG Arg	AAA Lys	TCC Ser	ATC Ile	2387
20	TAC Tyr 770	CGC	TAT Tyr	GAC Asp	CTG Leu	GCC Ala 775	TCG	GGA Gly	GCC Ala	ACC Thr	Glu	765 CAG Gln	TTG Leu	CCT Pro	CTC Leu	Thr	2435
	GGG	CTA Leu	CGG Arg	GCA Ala	GCA Ala 790	GTG	GCC Ala	CTG Leu	GAC Asp	TTT Phe 795	780 GAC Asp	TAT Tyr	GAG Glu	CAC His	Asn	785 TGT Cys	2483
25	TTG Leu	TAT Tyr	TGG Trp	TCC Ser 805	GAC	CTG Leu	GCC Ala	TTG Leu	GAC Asp 810	GTC	ATC Ile	CAG Gln	CGC Ar g	CTC Leu 815	800 TGT Cys	TTG Leu	2531
	AAT Asn	GGA Gly	AGC Ser 820	ACA Thr	GGG Gly	CAA Gln	GAG Glu	GTG Val 825	ATC	ATC Ile	AAT Asn	TCT Ser	GGC Gly 830	CTG	GAG Glu	ACA Thr	2579
30	GTA Val	GAA Glu 835	GCT Ala	TTG Leu	GCT Ala	TTT Phe	GAA Glu 840	CCC Pro	CTC Leu	AGC Ser	CAG Gln	CTG Leu 845	Curu	TAC Tyr	TGG Trp	GTA Val	2627
	GAT Asp 850	GCA Ala	GGC Gly	TTC Phe	AAA Lys	AAG Lys 855	ATT	GAG Glu	GTA Val	GCT Ala	AAT Asn 860	CCA	GAT Asp	GGC Gly	GAC Asp	Phe	2675
35	CGA Arg	CTC Leu	ACA Thr	ATC Ile	GTC Val 870	AAT	TCC Ser	TCT Ser	GTG Val	CTT Leu 875	GAT	CGT Ar g	CCC Pro	AGG Arg	GCT Ala 880	865 CTG Leu	2723
	GTC Val	CTC Leu	GTG Val	CCC Pro 885	CAA	GAG Glu	GGG Gly	GTG Val	ATG Met 890	TTC	TGG Trp	ACA Thr	GAC Asp	TGG Trp 895	CCA	GAC Asp	2771
40	ren	гÀг	900	GIY	Ile	Tyr	Arg	AGC Ser	AAT Asn	Met	Asp	Gly	Ser	GCT Ala	Ala	Tyr	2819
	nis	915	vaı	ser	GIU	Asp	920	AAG Lys	Trp	Pro	Asn	Gly 925	ATC Ile	Ser	Val	Asp	2867
45	930	GIN	Trp	ııe	Tyr	Trp 935	Thr	GAT Asp	Ala	Tyr	CTG Leu 940	GAG Glu	Cys	Ile	Glu	Arg	2915
	TIE	Thr	Phe	Ser	Gly 950	Gln	Gln	CGC Arg	Ser	Val 955	Ile	Leu	Asp	Asn	Leu 960	CCG Pro	2963
50	HIS	Pro	Tyr .	A1a 965	Ile .	Ala	Val	TTT Phe	Lys 970	Asn	Glu	Ile	Tyr	TGG : Trp :	GAT Asp	Asp	3011
	TGG Trp	TCA Ser	CAG Gln	CTC Leu	AGC Ser	ATA Ile	TTC Phe	CGA Arg	GCT Ala	TCC Ser	AAA Lys	TAC Tyr	AGT	GGG	TCC Ser	CAG Gln	3059

			980					985					940				
	ATG																3107
	Met	Glu	Ile	Leu	Ala	Asn	Gln	Leu	Thr	Gly	Leu	Met	Asp	Met	Lys	Ile	
5		995					1000					1005					
5				GGG													3155
		-	Lys	Gly	Lys			Gly	Ser	Asn			Val	Pro	Arg		
	1010					1015					1020				maa	1025	
				CTG													3203
	Cys	Ser	Leu	Leu	-		Pro	гла	Ala	1035		Ser	Arg	Sei	1040	_	
10	TOT	CCN	GNG	GAT	1030		AGC	AGT	GTG			TCA	GGG	GAC			3251
				Asp													3431
	Суз	FIO	GIU	1045		001			1050					1055			
	TGT	GAC	TGC	CCT		GGC	TAT	CAG			AAC	AAT	ACC			AAA	3299
	Cvs	Asp	Cys	Pro	Gln	Gly	Tyr	Gln	Leu	Lys	Asn	Asn	Thr	Cys	Val	Lys	
	_		1060)				1065	5				1070)			
15	GAA	GAG	AAC	ACC	TGT	CTT	CGC	AAC	CAG	TAT	CGC	TGC	AGC	AAC	GGG	AAC	3347
	Glu	Glu	Asn	Thr	Cys	Leu	Arg	Asn	Gln	Tyr	Arg			Asn	Gly	Asn	
		1075					1080					1085					
	TGT	ATC	AAC	AGC	ATT	TGG	TGG	TGT	GAC	TTT	GAC	AAC	GAC	TGT	GGA	GAC	3395
	-		Asn	Ser	Ile			Cys	Asp	Phe			Asp	Cys	GIY		
	1090					1095		~~		3.00	1100		030	CITIC	~~	1105	2442
20	ATG	AGC	GAT	GAG	AGA	AAC	TGC	CCT	ACC	ACC.	ATC	TGT	GAC	CIG	GAC	ACC The	3443
	Met	Ser	Asp	Glu			Сув	Pro	inr	1115		Cys	Asp	Leu	1120		
	CAC	TTT.	CCT	TGC	1110		ጥርጥ	ccc	ACT			CCA	CTG	TCC			3491
				Cys													3431
	GIII	Pile	Arg	1125		Gia	Ser	GLY	1130		110	110	LCu	1135		-70	
	тст	GAC	СТТ	GAG		GAC	тст	GGA			AGT	GAT	GAA			TGT	3539
<i>2</i> 5	Cvs	Asn	Leu	Glu	Asp	Asp	Cvs	Glv	Asp	Asn	Ser	Asp	Glu	Ser	His	Cvs	
	cys	p	1140				-1-	1149					115			4	
	GAA	ATG	CAC	CAG	TGC	CGG	AGT			TAC	AAC	TGC	AGT	TCC	GGC	ATG	3587
				Gln													
		1155	i		-	_	1160) _		_		-116	5				
	TGC	ATC	CGC	TCC	TCC	TGG	GTA	TGT	GAC	GGG	GAC	AAC	GAC	TGC	AGG	GAC	3635
30	Cys	Ile	Arg	Ser	Ser	Trp	Val	Суз	Asp	Gly	Asp	Asn	Asp	Cys	Arg	Asp	
	1170)				1175					118					1185	
	TGG	TCT	GAT	GAA	GCC	AAC	TGT	ACC	GCC	ATC	TAT	CAC	ACC	TGT	GAG	GCC	3683
	Trp	Ser	Asp	Glu			Cys	Thr	Ala			His	Thr	Cys			
	•		_		1190					119!					120		2721
	TCC	AAC	TTC	CAG	TGC	CGA	AAC	GGG	CAC	TGC	ATC	CCC	CAG	CGG	TGG	GCG	3731
35	Ser	Asn	Phe			Arg	Asn	GIA			TTE	Pro	GIN	121		Ala	
	man	~~~	~~~	120		CAC	mcc.	CAC	1210		TOO	CAT	CAG			GTC	3779
	TGT	GAC	GGG	Asp	ACG	BAC	Cuc	CAG	GAI	Cliv	Sex	GWI	GAG	AGN	Dro	Val	3113
	Cys	Asp	122		IIII	Asp	Cys	122		GIY	Ser	Asp	123		rio	Val	
	ልልሮ	тст			AAG	TGC	ТАА			CGC	TGC	CCA			ACT	TGC	3827
40	Asn	Cvs	Glu	Lvs	Lvs	Cvs	Asn	Glv	Phe	Arg	Cvs	Pro	Asn	Glv	Thr	Сув	
40	non	1239		L , 5	_, _	-7-	124			5	-1-	124		1		-2	
	ATC	CCA	TCC	AGC	AAA	CAT			GGT	CTG	CGT			TCT	GAT	GGC	3875
	Ile	Pro	Ser	Ser	Lys	His	Суз	Asp	Gly	Leu	Arg	Asp	Cys	Ser	Asp	Gly	
	125)			-	125	5				126	0				1265	
	TCC	GAT	GAA	CAG	CAC	TGC	GAG	CCC	CTC	TGT	ACG	CAC	TTC	ATG	GAC	TTT	3923
45	Ser	Asp	Glu	Gln	His	Cys	Glu	Pro	Leu	Cys	Thr	His	Phe	Met	Asp	Phe	
45					127	0				127					128		
	GTG	TGT	AAG	AAC	CGC	CAG	CAG	TGC	CTG	TTC	CAC	TCC	ATG	GTC	TGT	GAC	3971
	Val	Cys	Lys			Gln	Gln	Cys			His	Ser	Met	Val	_Cys	Asp	
				128				~~-	129			~~~	~~~	129		CCA	4010
	GGA	ATC	ATC	CAG	TGC	CGC	GAC	GGG	TCC	GAT	GAG	GAT	GCG	GCG	TTT	GCA	4019
50	GIY	ITe			Cys	Arg	Asp			Asp	GIU	Asp			rne	Ala	
30	003	maa	130	U (73.2	Cam		CAC	130	⊃ רארי	מא	ርጥአ	שיבית	131 CAT		ттс	GGT	4067
	GGA.	TGC	TCC	CAA G1⊷	DAT Acr	DYC	GAG Glii	Dhe	Hie	Lve	CIN	י בטר	Yen	Glii	Phe	Gly	
	GTÅ	131		GIU	veb	FIO	132		****	-ya	407	132				1	
		721	_				- J L	-					-				

						GGA											4115
	1330		Cys	GIN	ASII	1335		Cys	116	DET	1340		rp	ràs	cys	1345	
			GAT	GAT	TGC			TAT	тст	GAT			AAC	TGC	GAA		4163
5	Glv	Met	Asp	Asp	Cvs	Glv	Asp	Tyr	Ser	Asp	Glu	Ala	Asn	Cvs	Glu	Asn	4103
	1				1350			- 4 -		1355				-4-	1360		
	CCC	ACA	GAA	GCC	CCA	AAC	TGC	TCC	CGC	TAC	TTC	CAG	TTT	CGG	TGT	GAG	4211
	Pro	Thr	Glu	Ala	Pro	Asn	Cys	Ser			Phe	Gln	Phe			Glu	
				1365					1370					1375			
	AAT	GGC	CAC	TGC	ATC	CCC	AAC	AGA	TGG	AAA	TGT	GAC	AGG	GAG	AAC	GAC	4259
10	Asn	Gly			He	Pro	Asn	Arg		гÀг	Cys	Asp	Arg 1390		Asn	Asp	
	mam	aaa	1380		тст	CAT	CAC	1385 AAG		тст	CCA	CAT			ልጥጥ	Стт	4307
	Cve	GGG	Asn	Trn	Ser	ASD	Glu	Lys	Asp	Cvs	Glv	Asp	Ser	His	Ile	Leu	4507
	Cys	1395				11.55	1400			-,-	,	1405					
	CCC	TTC	TCG	ACT	CCT	GGG		TCC	ACG	TGT	CTG	CCC	AAT	TAC	TAC	CGC	4355
15								Ser									
	1410)				1415	5				1420					1425	
	TGC	AGC	AGT	GGG	ACC	TGC	GTG	ATG	GAC	ACC	TGG	GTG	TGC	GAC	GGG	TAC	4403
	Cys	Ser	Ser	Gly			Val	Met	Asp			Val	Cys	Asp			
	~~~	~ . m		~~1	1430		m c m	C3.C	CRC	1435		TCC	ccc	TTTC	1440		4451
	CGA	GAT	Cuc	Ala	GAT	Clar	Cor	GAC Asp	GAG	GAA	Δla	Cve	Pro	T.A.	LAU	Δla	443I
20	Arg	Asp	Cys	1445		Gry	261	rsp	1450		n.a	Cys	110	1455		AIG	
	AAC	GTC	ACT			TCC	ACT	CCC			CTT	GGG	CGA			CGA	4499
	Asn	Val	Thr	Ala	Ala	Ser	Thr	Pro	Thr	Gln	Leu	Gly	Arg	Сув	qaA	Arq	
			1460	)				1469	5				1470	)			
	TTT	GAG	TTC	GAA	TGC	CAC	CAA	CCG	AAG	ACG	TGT	ATT	CCC	AAC	TGG	AAG	4547
	Phe	Glu	Phe	Glu	Cys	His	Gln	Pro	Lys	Thr	Cys			Asn	$\mathtt{Trp}$	Lys	
25		1475	5				1480					1485					
	CGC	TGT	GAC	GGC	CAC	CAA	GAT	TGC	CAG	GAT	GGC	CGG	GAC	GAG	GCC	AAT	4595
			Asp	GIY	Hıs			Cys	GIn	Asp			Asp	GIU	Ата		
	1490	,	202	CAC	300	1495		ACT	TGC	እጥር	1500		GAG	ттс	CAG	1505 TGC	4643
	Cro	Dro	The	LAC	AGC	Thr	Len	Thr	Cve	Met	Ser	Ara	GAG	Phe	Gln	Cvs	4042
30	Cys	FIO	1111	птэ	1510		Beu	1111	Cys	151		149			1520		
	GAG	GAC	GGG	GAG	GCC	TGC	ATT	GTG	CTC			CGC	TGC	GAC			4691
	Glu	Asp	Gly	Glu	Ala	Cys	Ile	Val	Leu	Ser	Glu	Arg	Cys	Asp	Gly	Phe	
•		_	_	1525	5				1530	0				153	5		
	CTG	GAC	TGC	TCG	GAC	GAG	AGC	GAT	GAA	AAG	GCC	TGC	AGT	GAT	GAG	TTG	4739
	Leu	Asp			Asp	Glu	Ser	Asp		Lys	Ala	Cys			Glu	Leu	
35			1540	)				154		maa			1550		mær.	ccc	4707
	ACT	GTG	TAC	AAA	GTA	CAG	AAT	CTT	CAG	TGG	ACA	Ala	ACC	Dhe	Ser	Glv	4787
	Thr	1555		гÀг	vai	GIR	156	Leu	GIII	пр	1111	156		FILE	361	Gry	
	GAT	GTG.	ACTT	ጥፕር	ACC	TGG		AGG	CCC	AAA	AAA			TCT	GCA	TCT	4835
	Asp	Val	Thr	Leu	Thr	Trp	Met	Arg	Pro	Lvs	Lvs	Met	Pro	Ser	Ala	Ser	
	1570					157				•	158					1585	
40	TGT	GTA	TAT	AAT	GTC	TAC	TAC	AGG	GTG	GTT	GGA	GAG	AGC	ATA	TGG	AAG	4883
	Cys	Val	Tyr	Asn	Val	Tyr	Tyr	Arg	Val	Val	Gly	Glu	Ser	Ile	$\mathtt{Trp}$	Lys	
					159	0				159					160		
	ACT	CTG	GAG	ACC	CAC	AGC	AAT	AAG	ACA	AAC	ACT	GTA	TTA	AAA	GTC	TTG	4931
	Thr	Leu	Glu			Ser	Asn	Lys			Thr	vai	Leu	Lys	-var	Leu	
45			~~m	160	5	mam	CNC	GTT	161		CAC	CTT	CNG	161.		AGC	4979
45	AAA	CCA D=0	GAT	ACC	ACG	TAI	Gla	Val	Tare	Val	Gln	Val	Gln	Cvs	Leu	Ser	4712
	nys	PLO	162		1111	TYL	GIII	162	<i>шув</i> 5	val	G11:	741	163	0		J <b></b>	
	AAG	GCA	CAC	סבע	ACC	ТАА	GAC	TTT	GTG	ACC	CTG	AGG			GAG	GGA	5027
	Lvs	Ala	His	Asn	Thr	Asn	Asp	Phe	Val	Thr	Leu	Arg	Thr	Pro	Glu	Gly	
		163	5				164	0				164	5				
50	TTG	CCA	GAT	GCC	CCT	CGA	AAT	CTC	CAG	CTG	TÇA	CTC	CCC	AGG	GAA	GCA	5075
	Leu	Pro	Asp	Ala	Pro	Arg	Asn	Leu	Gln	Leu	Ser	Leu	Pro	Arg	Glu	Ala	
	165			_	_	165					166		a		<b></b>	1665	E 2 2 2
	CAA	CCT	GTG	ልጥተ	GTA	GGC	CAC	TGG	GCT	CCT	CCC	ATC	CAC	ACC	CAT	GGC	5123

	Glu	Gly	Val	Ile	Val	Gly	His	Trp	Ala	Pro	Pro	Ile	His	Thr		Gly	
	CTC	ATC	CGT	GAG			GTA	GAA	TO	167	AGG	3 CM			160	0	
	Leu	Ile	Arg	Glu	Tur	Tla	Val	Glu	TAC	AGC Com	Arg	AGI	GGT	TCC	AAG	ATG	5171
5			5	168	5 - 7 -			Giu	169	v 261	Arg	Set	GIY			Met	
	TGG	GCC	TCC	CAG	AGG	GCT	GCT	AGT	AAC	்ராரா	ΔCA	CAA	እጥሮ	169	> >>>	TTA	5010
	Trp	Ala	Ser	Gln	Arq	Ala	Ala	Ser	Asn	Phe	Thr	Glu	Tle	LAG	AAC	IIA	5219
			170	U				170	5				171	o '			
	TTG	GTC	AAC	ACT	CTA	TAC	ACC	GTC	AGA	GTG	GCT	GCG	GTG	ACT	AGT	CGT	5267
	Leu	Val	Asn	Thr	Leu	Tyr	Thr	Val	Arq	Val	Ala	Ala	Val	Thr	Ser	Ara	3207
10		171	5				172	0				172	5			_	
	GGA	ATA	GGA	AAC	TGG	AGC	GAT	TCT	AAA	TCC	ATT	ACC	ACC	ATA	AAA	GGA	5315
	Gly	Ile	Gly	Asn	Trp	Ser	Asp	Ser	Lys	Ser	Ile	Thr	Thr	Ile	Lys	Gly	
	173					173					174	)			_	1745	
	AAA	GTG	ATC	CCA	CCA	CCA	GAT	ATC	CAC	ATT	GAC	AGC	TAT	GGT	GAA	AAT	5363
	гàг	vai	iie	Pro	Pro	Pro	Asp	Ile	His	Ile	Asp	Ser	Tyr	Gly	Glu	Asn	
15	mam	CTTA	NCC	mma	1750		100		~~~	1759	5				176	0	
	THI	LAN	AGC Sex	Dho	Th-	CIG	ACC	ATG	GAG	AGT	GAT	ATC	AAG	GTG	AAT	GGC	5411
	TAT	Dea	261	176	= 1111	reu	Inr	Met			Asp	iie	Lys			Gly	
	тат	GTG	GTG	AAC	بلسك	דדר	TGG	GCA	1770	GAC	ACC	CNC	220	177!	2		
	Tvr	Val	Val	Asn	Leu	Phe	Trn	Δla	Dho	Acr	Thr	LAC	AAG	CAA	GAG	AGG	5459
	-1-		1780	)	200	1	11p	178	- F116	ASP	1111	urs	1790		GIU	Arg	
20	AGA	ACT			TTC	CGA	GGA	AGC.	בדב	TTC	TCA	CAC	T/30	(1977) (1977)	ccc	יייעגג	EE 0.7
	Arq	Thr	Leu	Asn	Phe	Arg	Glv	Ser	Ile	Leu	Ser	His	Lvs	Val	Glv	AAI	5507
		1799	5				1800	0				1809	5		_		
	CTG	ACA	GCT	CAT	ACA	TCC	TAT	GAG	ATT	TCT	GCC	TGG	GCC	AAG	ACT	GAC	5555
	Leu	Thr	Ala	His	Thr	Ser	Tyr	Glu	Ile	Ser	Ala	Trp	Ala	Lvs	Thr	Asp	3333
25	1810	)				1819	5				1820	)		-		1825	
25	TTG	GGG	GAT	AGC	CCT	CTG	GCA	TTT	GAG	CAT	GTT	ATG	ACC	AGA	GGG	GTT	5603
	Leu	Gly	Asp	Ser	Pro	Leu	Ala	Phe	Glu	His	Val	Met	Thr	Arg	Gly	Val	
	~~~		~~~		1830					1835	5				1840	כ	
	CGC	CCA D	CCT	GCA	CCT	AGC	CTC	AAG	GCC	AAA	GCC	ATC	AAC	CAG	ACT	GCA	5651
	Arg	PIO	PIO	1845	Pro	Ser	Leu	гàг			Ala	Ile	Asn			Ala	
30	GTG	GAA	тст			N.C.C	CCC	ccc	1850) """	GTG	~~~		1855			
	Val	Glu	Cvs	Thr	Trn	Thr	Glv	Pro	720	AAI	Val	GIT	TAT	GGT	ATT	TTC	5699
			1860)		****	O+ y	1865	, mry	ASII	VAI	vai	1870		TTE	Pne	
	TAT	GCC			TTT	CTT	GAC	CTC	TAT	CGC	AAC	CCG	AAG	»AGC	ጥጥር	አ ርጥ	5747
	Tyr	Ala	Thr	Ser	Phe	Leu	Asp	Leu	Tvr	Ara	Asn	Pro	Lvs	Ser	Leu	Thr	3/4/
		T8/2	•				1880)				1889	5				
35	ACT	TCA	CTC	CAC	AAC	AAG	ACG	GTC	ATT	GTC	AGT	AAG	GAT	GAG	CAG	TAT	5795
	Thr	Ser	Leu	His	Asn	Lys	Thr	Val	Ile	Val	Ser	Lys	Asp	Glu	Gln	Tyr	
•	1890	l				1895	5				1900)				1905	
	TTG	TTT	CTG	GTC	CGT	GTA	GTG	GTA	CCC	TAC	CAG	GGG	CCA	TCC	TCT	GAC	5843
	ren	Pne	Leu	vaı	Arg	Val	Val	Val	Pro		Gln	Gly	Pro	Ser			
	ጥ አ ር	CTT	CTL	CTC	1910		3.000	-	~~~	1915			· .		1920)	
40	Tur	Val	CIA	Mal.	LIC	Mot	ATC	Dwa	GAC	AGC	AGG	CTT	CCA	CCC	CGT	CAC	5891
	- y -	Vai	vai	1925	: Lys	MEC	116	PIO	1930	ser	Arg	Leu	Pro			His	
	CTG	CAT	GTG			ACG	GGC	מממ	1930	, TCC	GTG	CTC	A TO	1935	, TOC	a	5020
	Leu	His	Val	Val	His	Thr	Glv	Lvs	Thr	Ser	Val	Ual	TIA	Luc	T	Clu	5939
			1940)				1945		001	Val	Val	1950		TIP	Giu	
	TCA	CCG			TCT	CCT	GAC			TTG	TTG	TAT	GCA	АТТ	GCA	GTC	5987
4 5	Ser	Pro	Tyr	Asp	Ser	Pro	Asp	Gln	Asp	Leu	Leu	Tvr	Ala	Ile	Ala	Val	330,
		1955	,				1960)				1965	;				
	AAA	GAT	CTC	ATA	AGA	AAG	ACT	GAC	AGG	AGC	TAC	AAA	GTA	AAA	TCC	CGT	6035
	Lys	Asp	Leu	Ile	Arg	Lys	Thr	Asp	Arg	Ser	Tyr	Lys	Val	Lys	Ser	Arg	
	1970					1975	;				1980	1				1985	
5 0	AAC	AGC	ACT	GTG	GAA	TAC	ACC	CTT	AAC	AAG	TTG	GAG	CCT	GGC	GGG	AAA	6083
50	ASD	ser	Inr	val	GIU	Tyr	Thr	Leu	Asn	Lys	Leu	Glu	Pro	Gly			
	TAC	ראר	አጥር	ን ጥጥ	1990		OTTO	000	220	1995	3.00		~~~	mc-	2000)	
	Tvr	Hie	Tla	Lla	Val	CUV.	Ten	GI	MAC	MATG	AGC Ser	AAA	GAT	TCC	AGC Ca-	ATA	6131
	-1-		116	115	7 41	3111	neu	GIA	usii	rie C	JEI	nys	vab	SEL	SEL	TTG	

				2005	5				2010					3015			
	AAA	ATT	ACC	ACA	GTT	TCA	TTA	TCA	GCA	CCJ,	GAT	GCC	TTA	AAA	ATC	Α'·A	6179
	Lvs	Ile	Thr	Thr	Val	Ser	Leu	Ser	Ala	Pro	Asp	Ala	Leu	Lys	Ile	Ile	
i	-		2020)				2025	5				2030)			
	ACA	GAA	AAT	GAT	CAT	GTT	CTT	CTG	TTT	TGG	AAA	AGC	CTG	GCT	TTA	AAG	6227
	Thr	Glu	Asn	Asp	His	Val	Leu	Leu	Phe	Trp	Lys	Ser	Leu	Ala	Leu	Lys	
		2039	5	_			2040)				2045	5				
	GAA	AAG	CAT	TTT	AAT	GAA	AGC	AGG	GGC	TAT	GAG	ATA	CAC	ATG	TTT	GAT	6275
	Glu	Lys	His	Phe	Asn	Glu	Ser	Arg	Gly	Tyr	Glu	Ile	His	Met	Phe	Asp	
10	2050	٦				2055	5				2060)				2065	
	AGT	GCC	ATG	AAT	ATC	ACA	GCT	TAC	CTT	GGG	AAT	ACT	ACT	GAC	AAT	TTC	6323
	Ser	Ala	Met	Asn	Ile	Thr	Ala	Tyr	Leu	Gly	Asn	Thr	Thr	Asp	Asn	Phe	
					2070)				207	5				2080)	435 -
	TTT	AAA	ATT	TCC	AAC	CTG	AAG	ATG	GGT	CAT	AAT	TAC	ACG	TTC	ACC	GTC	6371
15	Phe	Lys	Ile			Leu	Lys	Met	Gly	His	Asn	Tyr	Thr	Pne	-Tnr	vaı	
. •				208	5				2090				~~~	209		3.000	6419
	CAA	GCA	AGA	TGC	CTT	TTT	GGC	AAC	CAG	ATC	TGT	93	GAG	CCI	310	AIC	6419
	Gln	Ala			Leu	Phe	GIA	Asn	GIn	TTE	Cys	GIY	211	PIO	Ala	116	
			210	o			~~~	210	~~~	CC3	CAT	CCN			N.C.C.	CNG	6467
	CTG	CTG	TAC	GAT	GAG	CTG	GGG	TCT	GGT	GCA	GAT	Ala	101	Ala	Th~	Gla	0407
20	Leu			Asp	GIU	ьeu	GIY	Ser	GIA	Ala	Asp	212	Ser	AIA	1111	GIII	
		211	5	m.cm	3.00	C N ITT	212	GCT	COT	CTC	GTG			ATC	тπъ	ጥጥር	6515
	GCT	GCC	AGA	TCT	ACG	GAI	GII	Ala	Ala	Val	Val	Val	Pro	Tle	Len	Phe	0323
			Arg	ser	1111	213	vai 5	AIG	AIG	VAI	214	0				2145	
	213) 	CTC	CTC	AGC	CTC	്ദേദ	GTG	GGG	ттт			CTG	TAC	ACG	AAG	6563
25	Tou	TIA	Teu	Leu	Ser	T.e.II	GIV	Val	Glv	Phe	Ala	Ile	Leu	Tyr	Thr	Lys	
	Leu	116	neu	neu	215		Q-7		1	215	5			-	216	0 -	
	CAC	CGG	AGG.	СТС	CAG	AGC	AGC	TTC	ACC	GCC	TTC	GCC	AAC	AGC	CAC	TAC	6611
	Hie	Arg	Ara	Len	Gln	Ser	Ser	Phe	Thr	Ala	Phe	Ala	Asn	Ser	His	Tyr	
				216	5				217	0				217	5		
	AGC	TCC	AGG	CTG	GGG	TCC	GCA	ATC	TTC	TCC	TCT	GGG	GAT	GAC	CTG	GGG	6659
30	Ser	Ser	Ara	Leu	Glv	Ser	Ala	Ile	Phe	Ser	Ser	Gly	Asp	Asp	Leu	Gly	
			218	Ω				218	5				219	0			
	GAA	GAT	CAT	GAA	GAT	GCC	CCT	ATG	ATA	ACT	GGA	TTT	TCA	GAT	GAC	GTC	6707
	Glu	Asp	Asp	Glu	Asp	Ala	Pro	Met	Ile	Thr	Gly	Phe	Ser	Asp	Asp	Val	
		219	5				220	0				220	5				
35	CCC	ATG	GTG	ATA	GCC	TGA	AAGA	GCT	TTCC	TCAC	TA G	AAAC	CAAA	T GG	TGTA	AATA	6762
	Pro	Met	Val	Ile	Ala												
	221	0													~===	~~ ~ ~ ~ ~	6000
	TTT	TATT	TGA	TAAA	GATA	GT T	GATG	GTTT	A TT	'TTAA	AAGA	TGC	ACTT	TGA	GTTG	CAATAT	6822
	GTT	ATTT	TTA	TATG	GGCC	AA A											6843

Claims

40

50

55

1. DNA having a nucleotide sequence as shown by Sequence ID No. 1.

- 2. An LDL receptor analog protein having an amino acid sequence as shown by Sequence ID No. 2 and coded by the DNA of Claim 1.
- 3. DNA having a nucleotide sequence as shown by Sequence ID No. 5.
- 4. An LDL receptor analog protein having an amino acid sequence as shown by Sequence ID No. 6 and coded by the DNA of Claim 3.
- 5. A recombinant vector comprising DNA as shown by Sequence ID No. 1 or 5 and a replicable vector.
- 6. Transformant cells which harbor the recombinant vector of Claim 5.

	7.	A method for the production of an LDL receptor analog protein comprising the steps of culturing the transformants of Claim 6 and collecting a polypeptide produced in the culture.
5		
10		
70		
15		
20		
25		
30		
35		
40		
4 5		
50		
i <i>5</i>		
-		

Europäisches Patentamt

European Patent Office

Office européen des brevets



(11) EP 0 773 290 A3

(12)

EUROPEAN PATENT APPLICATION

(88) Date of publication A3: 08.07.1998 Bulletin 1998/28

(51) Int. Cl.⁶: **C12N 15/12**, C07K 14/705 // C12N15/70, C12N15/79

(43) Date of publication A2: 14.05.1997 Bulletin 1997/20

(21) Application number: 96116108.0

(22) Date of filing: 08.10.1996

(84) Designated Contracting States:

AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC

NL PT SE

(30) Priority: 09.10.1995 JP 261440/95 24.04.1996 JP 102451/96

(71) Applicant: KOWA COMPANY, LTD.
Naka-ku Nagoya-shi Aichi-ken (37)

(72) Inventors:

Saito, Yasushi
 Chiba-shi, Chiba (JP)

Iwasaki, Akio
 Tsuchiura-shi, Ibaraki (JP)

 Arai, Koichi Urawa-shi, Saitama (JP)

Yamazaki, Hiroyuki
 Higashimurayama-shi, Tokyo (JP)

(74) Representative:
Wächtershäuser, Günter, Prof. Dr.
Patentanwalt,
Tal 29
80331 München (DE)

(54) Novel LDL receptor analog protein and the gene coding therefor

(57) The present invention is drawn to the gene of a novel LDL receptor family receptor which participates in lipoprotein metabolism, a critical factor that triggers the onset of arteriosclerosis.

The invention provides DNA having a nucleotide sequence as shown by Sequence ID No. 1 or No.5 is disclosed as well as rabbit tissue or human tissue LDL receptor analog protein having an amino acid sequence of Sequence ID No. 2 or 6 coded by such DNA.



EUROPEAN SEARCH REPORT

Application Number EP 96 11 6108

<u> </u>	DOCUMENTS CONSIDER Citation of document with indic		Relevant	CLASSIFICATION OF THE
ategory	Citation of document with indic of relevant passage	S anon, where appropriate,	to claim	APPLICATION (Int.CI.6)
A	WO 94 10322 A (UNIVER * the whole document	SITY OF TEXAS)	1-7	C12N15/12 C07K14/705 //C12N15/70,
A	WO 95 13374 A (BAYLOR * the whole document	COLLEGE OF MEDICINE)	1-7	C12N15/79
Α	KOZARSKY K. ET AL.: of low density lipopy deficiency in the Wath hyperlipemic rabbit wadenoviruses" JOURNAL OF BIOLOGICAL vol. 269, no. 18, 6 l pages 13695-13702, X * the whole document	rotein receptor Lanabe heritable With recombinant L CHEMISTRY, May 1994, MD US, P002062181	1-7	
A	gene transfer in viv improvement of hyper density lipoprotein rabbits"	cholesterolemia in low receptor-deficient L CHEMISTRY, January 1995, MD US, 062183	ł	TECHNICAL FIELDS SEARCHED (Int.CI.6)
	The present search report has t	peen drawn up for all claims	-	
	Place of search	Date of completion of the search		Examiner
ā	BERLIN	16 April 1998		Panzica, G
ž λ:	CATEGORY OF CITED DOCUMENTS particularly relevant if taken alone particularly relevant if combined with anot document of the same category	T : theory or print E : earlier patent after the filing D : document cite L : document cite	document, but is date and in the applica and for other reas	dion
O FORM	technological background non-written disclosure intermediate document	& : member of the document	e same patent f	amily, corresponding